

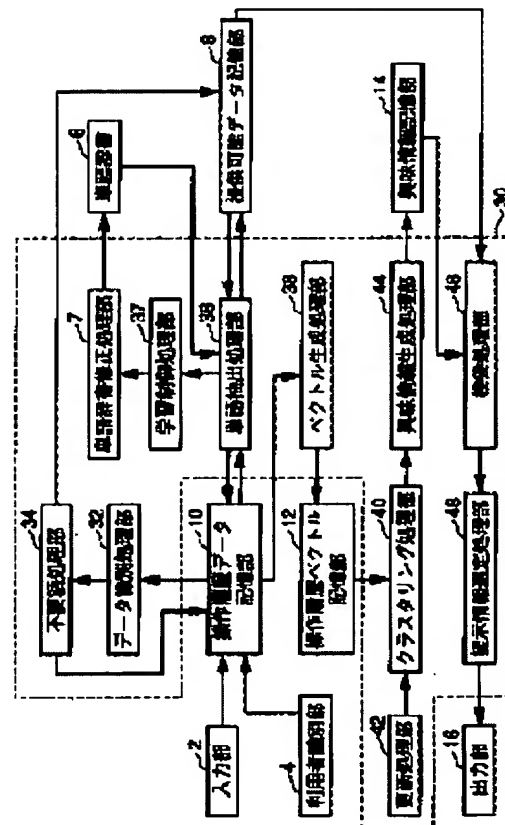
# DEVICE AND METHOD FOR INFORMATION PROVISION AND MACHINE-READABLE RECORDING MEDIUM WITH RECORDED PROGRAM MATERIALIZING THE SAME METHOD

**Patent number:** JP2001306612  
**Publication date:** 2001-11-02  
**Inventor:** KAWAMURA AKIYOSHI; KUROMUSHIYA KENICHI; KARASHI IKUO  
**Applicant:** SHARP CORP  
**Classification:**  
 - International: G06F17/30  
 - european:  
**Application number:** JP20000125711 20000426  
**Priority number(s):**

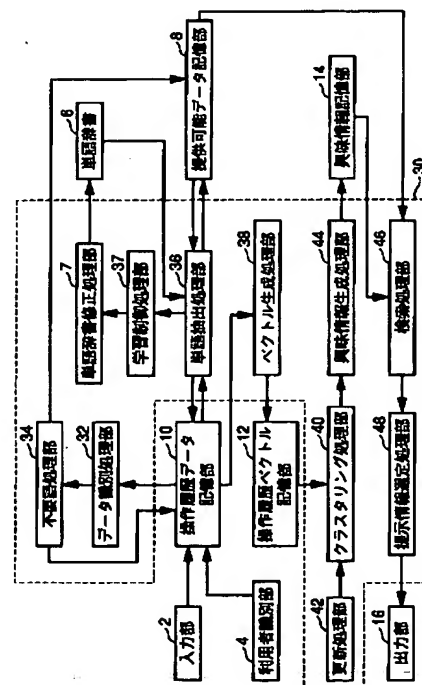
## Abstract of JP2001306612

**PROBLEM TO BE SOLVED:** To retrieve and provide information that a user takes a great interest in according to a use history.

**SOLUTION:** This device includes a providable data storage part 8 which stores providable information, an operation history data storage part 10 which stores use history information, a word dictionary 6 in which words and word feature vectors representing the features of the words are stored while they are made to correspond to each other, a word dictionary correction processing part 7 which corrects the word dictionary 6 according to the operation history data, an extraction processing part 36 which extracts words included in the providable information and use history information, a vector generation processing part 38 which generates feature vectors by pieces of use history information according to the word vectors corresponding to the extracted words, a clustering processing part 40 which classifies the use history information according to the distances of the feature vectors, and a retrieval processing part 46 which retrieves information that the user takes a great interest in from the providable information according to the classified use history information.



Data supplied from the esp@cenet database - Patent Abstracts of Japan



**【特許請求の範囲】**

【請求項1】 複数の単語と、各前記単語の特徴を表わす単語特徴ベクトルとを対応付けて記憶するための単語特徴記憶手段と、

利用者に対して提供可能な情報と、前記提供可能な情報の各々に対して前記単語特徴記憶手段に記憶された単語および単語特徴ベクトルに基づいて生成された提供可能情報特徴データとを準備するための提供可能情報準備手段と、

利用者が利用した情報の履歴情報を記憶するための利用履歴情報記憶手段と、

前記単語特徴記憶手段に記憶された単語および単語特徴ベクトルを参照し、前記利用履歴情報記憶手段に記憶された履歴情報を、各前記履歴情報に含まれる単語および対応の単語特徴ベクトルに基づいて複数のカテゴリに分類するための分類手段と、前記複数のカテゴリはそれぞれ、前記単語および単語特徴ベクトルに関連して定められる利用履歴情報特徴データにより特徴付けられ、前記単語特徴記憶手段に記憶された単語および単語特徴ベクトルを参照し、複数の前記履歴情報の各々についての特徴ベクトルを生成するための特徴ベクトル生成手段と、

予め定められた条件が満たされると、前記複数の履歴情報とその特徴ベクトルとに基づいて、前記単語特徴記憶手段の内容を更新するための単語特徴更新手段と、前記提供可能情報準備手段によって準備された提供可能な情報のうち、前記利用履歴情報特徴データと予め定められた関係にある提供可能情報特徴データを有するものを検索し出力するための提供情報検索手段とを含む、情報提供装置。

【請求項2】 前記単語特徴更新手段は、前記複数の履歴情報とその特徴ベクトルとに基づいて、前記単語特徴記憶手段に単語およびその単語に対応する単語特徴ベクトルを追加するための手段を含む、請求項1に記載の情報提供装置。

【請求項3】 前記単語特徴更新手段は、前記複数の履歴情報とその特徴ベクトルとに基づいて、前記単語特徴記憶手段に記憶された単語特徴ベクトルを更新するための手段を含む、請求項1または2に記載の情報提供装置。

【請求項4】 前記特徴ベクトル生成手段は、前記履歴情報に含まれる各単語に割当てられる第1の重みを算出するための第1の重み算出手段と、複数の前記履歴情報の各々について、前記履歴情報に含まれる単語に対応する単語特徴ベクトルに前記第1の重みを付けた第1のベクトルを算出し、算出された前記第1のベクトルを互いに加算するための加算手段とを含む、請求項1～3のいずれかに記載の情報提供装置。

【請求項5】 前記特徴ベクトル生成手段は、単語特徴ベクトルの成分ごとに割当てられる重みを算出するた

めの重み算出手段と、

複数の前記履歴情報の各々について、前記履歴情報に含まれる単語に対応する単語特徴ベクトルの成分に前記重みを付けたベクトルを算出し、算出された前記ベクトルを互いに加算するための加算手段とを含む、請求項1～3のいずれかに記載の情報提供装置。

【請求項6】 前記特徴ベクトル生成手段は、単語特徴ベクトルの成分ごとに割当てられる第2の重みを算出するための第2の重み算出手段をさらに含み、

前記加算手段は、複数の前記履歴情報の各々について、前記履歴情報に含まれる単語に対応する単語特徴ベクトルの成分に前記第2の重みを付けた第2のベクトルを算出し、前記第2のベクトルに前記第1の重みを付けた第1のベクトルを算出し、算出された前記第1のベクトルを互いに加算するための手段を含む、請求項4に記載の情報提供装置。

【請求項7】 前記情報提供装置は、利用者を識別するための利用者識別手段をさらに含み、

前記利用履歴記憶手段は、前記利用者別に情報を記憶するための複数の記憶手段を含み、

前記単語特徴更新手段は、前記利用者識別手段により識別された利用者が変更されたことを検知すると、単語特徴ベクトルを、前記識別された利用者の履歴情報とその特徴ベクトルとに基づいて、前記単語特徴記憶手段の内容を更新するための手段を含む、請求項1～6のいずれかに記載の情報提供装置。

【請求項8】 前記提供可能情報特徴データは、各提供可能な情報に出現する単語に対応する単語特徴ベクトルを含み、

前記利用履歴情報特徴データは、各分類ごとに含まれる利用履歴情報に出現する単語に対応する単語特徴ベクトルを含み、

前記提供情報検索手段は、前記提供可能情報特徴データに含まれる単語特徴ベクトルと前記利用履歴情報特徴データに含まれる単語特徴ベクトルとにより検索するための手段を含む、請求項1～7のいずれかに記載の情報提供装置。

【請求項9】 コンピュータを用いて利用者に情報を提供する方法であって、

複数の単語と、各前記単語の特徴を表わす単語特徴ベクトルとを対応付けて記憶する単語特徴記憶ステップと、

利用者に対して提供可能な情報と、前記提供可能な情報の各々に対して前記単語特徴記憶ステップにて記憶した単語および単語特徴ベクトルに基づいて生成された提供可能情報特徴データとを準備する提供可能情報準備ステップと、

利用者が利用した情報の履歴情報を記憶する利用履歴情報記憶ステップと、

前記単語特徴記憶ステップにて記憶した単語および単語

特徴ベクトルを参照し、前記利用履歴情報記憶ステップにて記憶した履歴情報を、各前記履歴情報に含まれる単語および対応の単語特徴ベクトルに基づいて複数のカテゴリに分類する分類ステップと、前記複数のカテゴリはそれぞれ、前記単語および単語特徴ベクトルに関連して定められる利用履歴情報特徴データにより特徴付けられ、

前記単語特徴記憶ステップにて記憶した単語および単語特徴ベクトルを参照し、複数の前記履歴情報の各々についての特徴ベクトルを生成する特徴ベクトル生成ステップと、

予め定められた条件が満たされると、前記複数の履歴情報とその特徴ベクトルとに基づいて、単語特徴記憶ステップにて記憶した内容を更新する単語特徴更新ステップと、

前記提供可能情報準備ステップにて準備した提供可能な情報のうち、前記利用履歴情報特徴データと予め定められた関係にある提供可能情報特徴データを有するものを検索し出力する提供情報検索ステップとを含む、情報提供方法。

【請求項10】 前記単語特徴更新ステップは、前記複数の履歴情報とその特徴ベクトルとに基づいて、前記単語特徴記憶ステップにて記憶した内容に単語およびその単語に対応する単語特徴ベクトルを追加するステップを含む、請求項9に記載の情報提供方法。

【請求項11】 前記単語特徴更新ステップは、前記複数の履歴情報とその特徴ベクトルとに基づいて、前記単語特徴記憶ステップにて記憶した単語特徴ベクトルを更新するステップを含む、請求項9または10に記載の情報提供方法。

【請求項12】 前記特徴ベクトル生成ステップは、前記履歴情報に含まれる各単語に割当てられる第1の重みを算出する第1の重み算出ステップと、複数の前記履歴情報の各々について、前記履歴情報に含まれる単語に対応する単語特徴ベクトルに前記第1の重みを付けた第1のベクトルを算出し、算出された前記第1のベクトルを互いに加算する加算ステップとを含む、請求項9～11のいずれかに記載の情報提供方法。

【請求項13】 前記特徴ベクトル生成ステップは、単語特徴ベクトルの成分ごとに割当てられる重みを算出する重み算出ステップと、

複数の前記履歴情報の各々について、前記履歴情報に含まれる単語に対応する単語特徴ベクトルの成分に前記重みを付けたベクトルを算出し、算出された前記ベクトルを互いに加算する加算ステップとを含む、請求項9～11のいずれかに記載の情報提供方法。

【請求項14】 前記特徴ベクトル生成ステップは、単語特徴ベクトルの成分ごとに割当てられる第2の重みを算出する第2の重み算出ステップをさらに含み、前記加算ステップは、複数の前記履歴情報の各々につい

て、前記履歴情報に含まれる単語に対応する単語特徴ベクトルの成分に前記第2の重みを付けた第2のベクトルを算出し、前記第2のベクトルに前記第1の重みを付けた第1のベクトルを算出し、算出された前記第1のベクトルを互いに加算するステップを含む、請求項12に記載の情報提供方法。

【請求項15】 前記情報提供方法は、利用者を識別する利用者識別ステップをさらに含み、

前記利用履歴記憶ステップは、前記利用者別に情報を記憶する複数の記憶ステップを含み、

前記単語特徴更新ステップは、前記利用者識別ステップにより識別された利用者の変更されたことを検知すると、単語特徴ベクトルを、前記識別された利用者の履歴情報とその特徴ベクトルとに基づいて、前記単語特徴記憶ステップにて記憶した内容を更新するステップを含む、請求項9～14のいずれかに記載の情報提供方法。

【請求項16】 前記提供可能情報特徴データは、各提供可能な情報に出現する単語に対応する単語特徴ベクトルを含み、

前記利用履歴情報特徴データは、各分類ごとに含まれる利用履歴情報に出現する単語に対応する単語特徴ベクトルを含み、

前記提供情報検索ステップは、前記提供可能情報特徴データに含まれる単語特徴ベクトルと前記利用履歴情報特徴データに含まれる単語特徴ベクトルとにより検索するステップを含む、請求項9～15のいずれかに記載の情報提供方法。

【請求項17】 コンピュータを用いて利用者に情報を提供する方法を実現するプログラムを記録した機械読取可能な記録媒体であって、

前記情報提供方法は、

複数の単語と、各前記単語の特徴を表わす単語特徴ベクトルとを対応付けて記憶する単語特徴記憶ステップと、

利用者に対して提供可能な情報と、前記提供可能な情報の各々に対して前記単語特徴記憶ステップにて記憶した単語および単語特徴ベクトルに基づいて生成された提供可能情報特徴データとを準備する提供可能情報準備ステップと、

利用者が利用した情報の履歴情報を記憶する利用履歴情報記憶ステップと、

前記単語特徴記憶ステップにて記憶した単語および単語特徴ベクトルを参照し、前記利用履歴情報記憶ステップにて記憶した履歴情報を、各前記履歴情報に含まれる単語および対応の単語特徴ベクトルに基づいて複数のカテゴリに分類する分類ステップと、前記複数のカテゴリはそれぞれ、前記単語および単語特徴ベクトルに関連して定められる利用履歴情報特徴データにより特徴付けられ、

前記単語特徴記憶ステップにて記憶した単語および単語

特徴ベクトルを参照し、複数の前記履歴情報の各々についての特徴ベクトルを生成する特徴ベクトル生成ステップと、

予め定められた条件が満たされると、前記複数の履歴情報とその特徴ベクトルとに基づいて、単語特徴記憶ステップにて記憶した内容を更新する単語特徴更新ステップと、

前記提供可能情報準備ステップにて準備した提供可能な情報のうち、前記利用履歴情報特徴データと予め定められた関係にある提供可能情報特徴データを有するものを検索し出力する提供情報検索ステップとを含む、情報提供方法を実現するプログラムを記録した機械読取可能な記録媒体。

【請求項18】 前記単語特徴更新ステップは、前記複数の履歴情報とその特徴ベクトルとに基づいて、前記単語特徴記憶ステップにて記憶した内容に単語およびその単語に対応する単語特徴ベクトルを追加するステップを含む、請求項17に記載の機械読取可能な記録媒体。

【請求項19】 前記単語特徴更新ステップは、前記複数の履歴情報とその特徴ベクトルとに基づいて、前記単語特徴記憶ステップにて記憶した単語特徴ベクトルを更新するステップを含む、請求項17または18に記載の機械読取可能な記録媒体。

【請求項20】 前記特徴ベクトル生成ステップは、前記履歴情報に含まれる各単語に割当てられる第1の重みを算出する第1の重み算出ステップと、

複数の前記履歴情報の各々について、前記履歴情報に含まれる単語に対応する単語特徴ベクトルに前記第1の重みを付けた第1のベクトルを算出し、算出された前記第1のベクトルを互いに加算する加算ステップとを含む、請求項17～19のいずれかに記載の機械読取可能な記録媒体。

【請求項21】 前記特徴ベクトル生成ステップは、単語特徴ベクトルの成分ごとに割当てられる重みを算出する重み算出ステップと、

複数の前記履歴情報の各々について、前記履歴情報に含まれる単語に対応する単語特徴ベクトルの成分に前記重みを付けたベクトルを算出し、算出された前記ベクトルを互いに加算する加算ステップとを含む、請求項17～19のいずれかに記載の機械読取可能な記録媒体。

【請求項22】 前記特徴ベクトル生成ステップは、単語特徴ベクトルの成分ごとに割当てられる第2の重みを算出する第2の重み算出ステップをさらに含み、前記加算ステップは、複数の前記履歴情報の各々について、前記履歴情報に含まれる単語に対応する単語特徴ベクトルの成分に前記第2の重みを付けた第2のベクトルを算出し、前記第2のベクトルに前記第1の重みを付けた第1のベクトルを算出し、算出された前記第1のベクトルを互いに加算するステップを含む、請求項20に記載の機械読取可能な記録媒体。

【請求項23】 前記情報提供方法は、利用者を識別する利用者識別ステップをさらに含み、

前記利用履歴記憶ステップは、前記利用者別に情報を記憶する複数の記憶ステップを含み、

前記単語特徴更新ステップは、前記利用者識別ステップにより識別された利用者に変更されたことを検知すると、単語特徴ベクトルを、前記識別された利用者の履歴情報とその特徴ベクトルとに基づいて、前記単語特徴記憶ステップにて記憶した内容を更新するステップを含む、請求項17～22のいずれかに記載の機械読取可能な記録媒体。

【請求項24】 前記提供可能情報特徴データは、各提供可能な情報に出現する単語に対応する単語特徴ベクトルを含み、

前記利用履歴情報特徴データは、各分類ごとに含まれる利用履歴情報に出現する単語に対応する単語特徴ベクトルを含み、

前記提供情報検索ステップは、前記提供可能情報特徴データに含まれる単語特徴ベクトルと前記利用履歴情報特徴データに含まれる単語特徴ベクトルとにより検索するステップを含む、請求項17～23のいずれかに記載の機械読取可能な記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、利用者の操作履歴から利用者の興味の高い情報を検索し提供する装置に関し、特に、利用者の操作履歴を分析した後、その分析結果に基づいて利用者にとって興味の高い情報を検索する装置に関する。

【0002】

【従来の技術】従来の情報提供技術において、利用者が興味のある情報の提供を受けるには、予め関連するキーワードを利用者自身の手で登録しておかなければならない。

【0003】このようなキーワードの登録を避けるものとして、特開平7-56929号公報に開示された履歴利用データベース検索方式がある。この発明は、利用者がかつて購入した物件の履歴を保持する購入履歴データベースと、購入対象データを保持する物件データベースと、登録データと利用者がかつて購入した物件との類似度を計算する距離計算部と、登録物件を類似度の降順に整列する類似物件整列部と、類似物件を推奨物件として表示する推奨物件表示部とを含む。この発明によると、過去の購入履歴を参照し、利用者がかつて購入した物件と類似した傾向を有する物件を検索することができる。そして次の物件購入の際に、利用者はその検索結果を推奨物件として参考にして、購入の意思決定を行なうことができる。

【0004】

【発明が解決しようとする課題】しかし、前述の発明で

は、物件データベースのデータ量および購入履歴データベースのデータ量が大きくなった場合、購入履歴データベース内の個々の履歴データと、物件データベース内の物件データとの類似度を個々に計算し、検索していたのでは多大な時間がかかる。また、前述の発明は、購入履歴のうちの一部に検索対象を制限するための付加情報を入力し、検索回数を減らす処理も含むが、選択された購入履歴だけで利用者の興味を完全に表現する検索はできない。

【0005】そこで、本発明は前述の課題を解決するためになされたものであって、その目的は、利用者の利用履歴から利用者の興味の高い情報を迅速かつ的確に抽出し、その情報を用いて利用者の興味を反映した情報を検索し提供することができる情報提供装置、情報提供方法、およびそれらを実現するプログラムを記録した機械読取可能な記録媒体を提供することである。

【0006】

【課題を解決するための手段】第1の発明に係る情報提供装置は、複数の単語と、各単語の特徴を表わす単語特徴ベクトルとを対応付けて記憶するための単語特徴記憶手段と、利用者に対して提供可能な情報と、提供可能な情報の各々に対して単語特徴記憶手段に記憶された単語および単語特徴ベクトルに基づいて生成された提供可能情報特徴データとを準備するための提供可能情報準備手段と、利用者が利用した情報の履歴情報を記憶するための利用履歴情報記憶手段と、単語特徴記憶手段に記憶された単語および単語特徴ベクトルを参照し、利用履歴情報記憶手段に記憶された履歴情報を、各履歴情報に含まれる単語および対応の単語特徴ベクトルに基づいて複数のカテゴリに分類するための分類手段と、複数のカテゴリはそれぞれ、単語および単語特徴ベクトルに関連して定められる利用履歴情報特徴データにより特徴付けられ、単語特徴記憶手段に記憶された単語および単語特徴ベクトルを参照し、複数の履歴情報の各々についての特徴ベクトルを生成するための特徴ベクトル生成手段と、予め定められた条件が満たされると、複数の履歴情報とその特徴ベクトルとに基づいて、単語特徴記憶手段の内容を更新するための単語特徴更新手段と、提供可能情報準備手段によって準備された提供可能な情報のうち、利用履歴情報特徴データと予め定められた関係にある提供可能情報特徴データを有するものを検索し出力するための提供情報検索手段とを含む。

【0007】第1の発明によると、単語特徴記憶手段は、複数の単語と、各単語の特徴を表わす単語特徴ベクトルとを対応付けて記憶する。提供可能情報準備手段は、利用者に対して提供可能な情報と、提供可能な情報の各々に対して単語特徴記憶手段に記憶された単語および単語特徴ベクトルに基づいて生成された提供可能情報特徴データとを準備する。利用履歴情報記憶手段は、利用者が利用した情報の履歴情報を記憶する。分類手段

は、単語特徴記憶手段に記憶された単語および単語特徴ベクトルを参照し、利用履歴情報記憶手段に記憶された履歴情報を、各履歴情報に含まれる単語および対応の単語特徴ベクトルに基づいて複数のカテゴリに分類する。特徴ベクトル生成手段は、単語特徴記憶手段に記憶された単語および単語特徴ベクトルを参照し、複数の履歴情報の各々についての特徴ベクトルを生成する。単語特徴更新手段は、予め定められた条件が満たされると、複数の履歴情報とその特徴ベクトルとに基づいて、単語特徴記憶手段の内容を更新する。提供情報検索手段は、提供可能情報準備手段によって準備された提供可能な情報のうち、利用履歴情報特徴データと予め定められた関係にある提供可能情報特徴データを有するものを検索し出力する。これにより、単語特徴記憶手段の内容は逐次更新され、最新の利用者の履歴情報に基づく単語およびその単語の単語特徴ベクトルとが記憶される。これに加えて、利用履歴情報自体ではなく単語特徴記憶手段に記憶された内容に基づき分類されたカテゴリごとの特徴データに従って提供情報を検索することができるため、利用者に提供すべき情報を迅速かつ的確に検索できる。その結果、利用者の利用履歴から利用者の興味の高い情報を迅速かつ的確に抽出し、その情報を用いて利用者の興味を反映した情報を検索し提供することができる情報提供装置を提供することができる。

【0008】第2の発明に係る情報提供装置は、第1の発明の構成に加えて、単語特徴更新手段は、複数の履歴情報とその特徴ベクトルとに基づいて、単語特徴記憶手段に単語およびその単語に対応する単語特徴ベクトルを追加するための手段を含む。

【0009】第2の発明によると、単語特徴更新手段は、複数の履歴情報とその特徴ベクトルとに基づいて、単語特徴記憶手段に単語およびその単語に対応する単語特徴ベクトルを追加する。これにより、単語特徴記憶手段に記憶されていなかった単語およびその単語の単語特徴ベクトルを生成し、記憶させることができる。その結果、当初記憶させた単語およびその単語の単語特徴ベクトルに加えて、多数の単語に基づいて、利用者の利用履歴から利用者の興味の高い情報を迅速かつ的確に抽出し、その情報を用いて利用者の興味を反映した情報を検索し提供することができる情報提供装置を提供することができる。

【0010】第3の発明に係る情報提供装置は、第1または2の発明の構成に加えて、単語特徴更新手段は、複数の履歴情報とその特徴ベクトルとに基づいて、単語特徴記憶手段に記憶された単語特徴ベクトルを更新するための手段を含む。

【0011】第3の発明によると、単語特徴更新手段は、複数の履歴情報とその特徴ベクトルとに基づいて、単語特徴記憶手段に記憶された単語特徴ベクトルを更新する。これにより、利用履歴情報の変化に対応させて単



語特徴ベクトルを生成し、単語特徴記憶手段に記憶させることができる。その結果、利用履歴情報の変化に対応させて当初記憶させた単語特徴ベクトルを更新して、更新された単語特徴ベクトルに基づいて、利用者の利用履歴から利用者の興味の高い情報を迅速かつ的確に抽出し、その情報を用いて利用者の興味を反映した情報を検索し提供することができる情報提供装置を提供することができる。

【0012】第4の発明に係る情報提供装置は、第1～3のいずれかの発明の構成に加えて、特徴ベクトル生成手段は、履歴情報に含まれる各単語に割当てられる第1の重みを算出するための第1の重み算出手段と、複数の履歴情報の各々について、履歴情報に含まれる単語に対応する単語特徴ベクトルに第1の重みを付けた第1のベクトルを算出し、算出された第1のベクトルを互いに加算するための加算手段とを含む。

【0013】第4の発明によると、第1の重み算出手段は、履歴情報に含まれる各単語に割当てられる第1の重みを算出する。加算手段は、複数の履歴情報の各々について、履歴情報に含まれる単語に対応する単語特徴ベクトルに第1の重みを付けた第1のベクトルを算出し、算出された第1のベクトルを互いに加算する。これにより、利用者の履歴情報ごとの特徴ベクトルを算出する場合において、単語に対応させた第1の重み（たとえば、単語の出現頻度に対応させて設定する重み）に基づいて履歴情報ごとの特徴ベクトルを算出できる。その結果、単語特徴記憶手段に記憶された単語特徴ベクトルを適切に更新することができる。

【0014】第5の発明に係る情報提供装置は、第1～3のいずれかの発明の構成に加えて、特徴ベクトル生成手段は、単語特徴ベクトルの成分ごとに割当てられる重みを算出するための重み算出手段と、複数の履歴情報の各々について、履歴情報に含まれる単語に対応する単語特徴ベクトルの成分に重みを付けたベクトルを算出し、算出されたベクトルを互いに加算するための加算手段とを含む。

【0015】第5の発明によると、重み算出手段は、単語特徴ベクトルの成分ごとに割当てられる重みを算出する。加算手段は、複数の履歴情報の各々について、履歴情報に含まれる単語に対応する単語特徴ベクトルの成分に重みを付けたベクトルを算出し、算出されたベクトルを互いに加算する。これにより、利用者の履歴情報ごとの特徴ベクトルを算出する場合において、単語ベクトルの成分に対応させた重み（たとえば、単語ベクトルの特定の成分の出現頻度に対応させて設定する重み）に基づいて履歴情報ごとの特徴ベクトルを算出できる。その結果、単語特徴記憶手段に記憶された単語特徴ベクトルを適切に更新することができる。

【0016】第6の発明に係る情報提供装置は、第4の発明の構成に加えて、特徴ベクトル生成手段は、単語特

徴ベクトルの成分ごとに割当てられる第2の重みを算出するための第2の重み算出手段をさらに含み、加算手段は、複数の履歴情報の各々について、履歴情報に含まれる単語に対応する単語特徴ベクトルの成分に第2の重みを付けた第2のベクトルを算出し、第2のベクトルに第1の重みを付けた第1のベクトルを算出し、算出された第1のベクトルを互いに加算するための手段を含む。

【0017】第6の発明によると、第2の重み算出手段は、単語特徴ベクトルの成分ごとに割当てられる第2の重みを算出する。加算手段は、複数の履歴情報の各々について、履歴情報に含まれる単語に対応する単語特徴ベクトルの成分に第2の重みを付けた第2のベクトルを算出し、第2のベクトルに第1の重みを付けた第1のベクトルを算出し、算出された第1のベクトルを互いに加算する。これにより、利用者の履歴情報ごとの特徴ベクトルを算出する場合において、単語ベクトルの成分に対応させた第2の重み（たとえば、単語ベクトルの特定の成分の出現頻度に対応させて設定する重み）と、単語に対応させた第1の重み（たとえば、単語の出現頻度に対応させて設定する重み）とに基づいて履歴情報ごとの特徴ベクトルを算出できる。その結果、単語特徴記憶手段に記憶された単語特徴ベクトルを適切に更新することができる。

【0018】第7の発明に係る情報提供装置は、第1～6のいずれかの発明の構成に加えて、利用者を識別するための利用者識別手段をさらに含み、利用履歴記憶手段は、利用者別に情報を記憶するための複数の記憶手段を含み、単語特徴更新手段は、利用者識別手段により識別された利用者が変更されたことを検知すると、単語特徴ベクトルを、識別された利用者の履歴情報とその特徴ベクトルとに基づいて、単語特徴記憶手段の内容を更新するための手段を含む。

【0019】第7の発明によると、利用者識別手段は利用者を識別し、利用履歴記憶手段は、利用者別に情報を記憶する。単語特徴更新手段は、利用者識別手段により識別された利用者が変更されたことを検知すると、単語特徴ベクトルを、識別された利用者の履歴情報とその特徴ベクトルとに基づいて、単語特徴記憶手段の内容を更新する。これにより、利用者ごとの履歴情報に基づいて、単語特徴記憶手段の内容を更新することができる。その結果、利用者ごとに利用者の利用履歴から利用者の興味の高い情報を迅速かつ的確に抽出し、その情報を用いて利用者の興味を反映した情報を検索し提供することができる情報提供装置を提供することができる。

【0020】第8の発明に係る情報提供装置は、第1～7のいずれかの発明の構成に加えて、提供可能情報特徴データは、各提供可能な情報に出現する単語に対応する単語特徴ベクトルを含み、利用履歴情報特徴データは、各分類ごとに含まれる利用履歴情報に出現する単語に対応する単語特徴ベクトルを含み、提供情報検索手段は、

提供可能情報特徴データに含まれる単語特徴ベクトルと利用履歴情報特徴データに含まれる単語特徴ベクトルとにより検索するための手段を含む。

【0021】第8の発明によると、提供情報検索手段は、各提供情報に出現する単語特徴ベクトルと各カテゴリに含まれる利用履歴情報に出現する単語に対応した単語特徴ベクトルに基づいて利用者の興味が高い情報を検索する。これにより、ベクトルによる演算を用いて、情報を検索することができ、利用者に提供する情報をさらに正確に検索できる情報提供装置を実現できる。その結果、利用者の利用履歴から利用者の興味の高い情報を迅速かつ的確に抽出し、その情報を用いて利用者の興味を反映した情報を検索し提供することができる情報提供装置を提供することができる。

【0022】第9の発明に係る情報提供方法は、コンピュータを用いて利用者に情報を提供する方法であって、複数の単語と、各単語の特徴を表わす単語特徴ベクトルとを対応付けて記憶する単語特徴記憶ステップと、利用者に対して提供可能な情報と、提供可能な情報の各々に対して単語特徴記憶ステップにて記憶した単語および単語特徴ベクトルに基づいて生成された提供可能情報特徴データとを準備する提供可能情報準備ステップと、利用者が利用した情報の履歴情報を記憶する利用履歴情報記憶ステップと、単語特徴記憶ステップにて記憶した単語および単語特徴ベクトルを参照し、利用履歴情報記憶ステップにて記憶した履歴情報を、各履歴情報に含まれる単語および対応の単語特徴ベクトルに基づいて複数のカテゴリに分類する分類ステップと、複数のカテゴリはそれぞれ、単語および単語特徴ベクトルに関連して定められる利用履歴情報特徴データにより特徴付けられ、単語特徴記憶ステップにて記憶した単語および単語特徴ベクトルを参照し、複数の履歴情報の各々についての特徴ベクトルを生成する特徴ベクトル生成ステップと、予め定められた条件が満たされると、複数の履歴情報とその特徴ベクトルとに基づいて、単語特徴記憶ステップにて記憶した内容を更新する単語特徴更新ステップと、提供可能情報準備ステップにて準備した提供可能な情報のうち、利用履歴情報特徴データと予め定められた関係にある提供可能情報特徴データを有するものを検索し出力する提供情報検索ステップとを含む。

【0023】第9の発明によると、単語特徴記憶ステップは、複数の単語と、各単語の特徴を表わす単語特徴ベクトルとを対応付けて記憶する。提供可能情報準備ステップは、利用者に対して提供可能な情報と、提供可能な情報の各々に対して単語特徴記憶ステップにて記憶した単語および単語特徴ベクトルに基づいて生成された提供可能情報特徴データとを準備する。利用履歴情報記憶ステップは、利用者が利用した情報の履歴情報を記憶する。分類ステップは、単語特徴記憶ステップにて記憶した単語および単語特徴ベクトルを参照し、利用履歴情報

記憶ステップにて記憶した履歴情報を、各履歴情報に含まれる単語および対応の単語特徴ベクトルに基づいて複数のカテゴリに分類する。特徴ベクトル生成ステップは、単語特徴記憶ステップにて記憶した単語および単語特徴ベクトルを参照し、複数の履歴情報の各々についての特徴ベクトルを生成する。単語特徴更新ステップは、予め定められた条件が満たされると、複数の履歴情報とその特徴ベクトルとに基づいて、単語特徴記憶ステップにて記憶した内容を更新する。提供情報検索ステップは、提供可能情報準備ステップにて準備した提供可能な情報のうち、利用履歴情報特徴データと予め定められた関係にある提供可能情報特徴データを有するものを検索し出力する。これにより、単語特徴記憶ステップにより記憶される内容は逐次更新され、最新の利用者の履歴情報に基づく単語およびその単語の単語特徴ベクトルとが記憶される。これに加えて、利用履歴情報自体ではなく単語特徴記憶ステップにて記憶した内容に基づき分類されたカテゴリごとの特徴データに従って提供情報を検索することができるため、利用者に提供すべき情報を迅速かつ的確に検索できる。その結果、利用者の利用履歴から利用者の興味の高い情報を迅速かつ的確に抽出し、その情報を用いて利用者の興味を反映した情報を検索し提供することができる情報提供方法を提供することができる。

【0024】第10の発明に係る情報提供方法は、第9の発明の構成に加えて、単語特徴更新ステップは、複数の履歴情報とその特徴ベクトルとに基づいて、単語特徴記憶ステップにて記憶した内容に単語およびその単語に対応する単語特徴ベクトルを追加するステップを含む。

【0025】第10の発明によると、単語特徴更新ステップは、複数の履歴情報とその特徴ベクトルとに基づいて、単語特徴記憶ステップにて記憶した内容に単語およびその単語に対応する単語特徴ベクトルを追加する。これにより、単語特徴記憶ステップにて記憶されていなかった単語およびその単語の単語特徴ベクトルを生成し、記憶させることができる。その結果、当初記憶させた単語およびその単語の単語特徴ベクトルに加えて、多数の単語に基づいて、利用者の利用履歴から利用者の興味の高い情報を迅速かつ的確に抽出し、その情報を用いて利用者の興味を反映した情報を検索し提供することができる情報提供方法を提供することができる。

【0026】第11の発明に係る情報提供方法は、第9または10の発明の構成に加えて、単語特徴更新ステップは、複数の履歴情報とその特徴ベクトルとに基づいて、単語特徴記憶ステップにて記憶した単語特徴ベクトルを更新するステップを含む。

【0027】第11の発明によると、単語特徴更新ステップは、複数の履歴情報とその特徴ベクトルとに基づいて、単語特徴記憶ステップにて記憶した単語特徴ベクトルを更新する。これにより、利用履歴情報の変化に対応



させて単語特徴ベクトルを生成し、単語特徴更新ステップにて更新することができる。その結果、利用履歴情報の変化に対応させて当初記憶させた単語特徴ベクトルを更新して、更新された単語特徴ベクトルに基づいて、利用者の利用履歴から利用者の興味の高い情報を迅速かつ的確に抽出し、その情報を用いて利用者の興味を反映した情報を検索し提供することができる情報提供方法を提供することができる。

【0028】第12の発明に係る情報提供方法は、第9～11のいずれかの発明の構成に加えて、特徴ベクトル生成ステップは、履歴情報に含まれる各単語に割当てられる第1の重みを算出する第1の重み算出ステップと、複数の履歴情報の各々について、履歴情報に含まれる単語に対応する単語特徴ベクトルに第1の重みを付けた第1のベクトルを算出し、算出された第1のベクトルを互いに加算する加算ステップとを含む。

【0029】第12の発明によると、第1の重み算出ステップは、履歴情報に含まれる各単語に割当てられる第1の重みを算出する。加算ステップは、複数の履歴情報の各々について、履歴情報に含まれる単語に対応する単語特徴ベクトルに第1の重みを付けた第1のベクトルを算出し、算出された第1のベクトルを互いに加算する。これにより、利用者の履歴情報ごとの特徴ベクトルを算出する場合において、単語に対応させた第1の重み（たとえば、単語の出現頻度に対応させて設定する重み）に基づいて履歴情報ごとの特徴ベクトルを算出できる。その結果、単語特徴記憶ステップにて記憶した単語特徴ベクトルを適切に更新することができる。

【0030】第13の発明に係る情報提供方法は、第9～11のいずれかの発明の構成に加えて、特徴ベクトル生成ステップは、単語特徴ベクトルの成分ごとに割当てられる重みを算出する重み算出ステップと、複数の履歴情報の各々について、履歴情報に含まれる単語に対応する単語特徴ベクトルの成分に重みを付けたベクトルを算出し、算出されたベクトルを互いに加算する加算ステップとを含む。

【0031】第13の発明によると、重み算出ステップは、単語特徴ベクトルの成分ごとに割当てられる重みを算出する。加算ステップは、複数の履歴情報の各々について、履歴情報に含まれる単語に対応する単語特徴ベクトルの成分に重みを付けたベクトルを算出し、算出されたベクトルを互いに加算する。これにより、利用者の履歴情報ごとの特徴ベクトルを算出する場合において、単語ベクトルの成分に対応させた重み（たとえば、単語ベクトルの特定の成分の出現頻度に対応させて設定する重み）に基づいて履歴情報ごとの特徴ベクトルを算出できる。その結果、単語特徴記憶ステップにて記憶した単語特徴ベクトルを適切に更新することができる。

【0032】第14の発明に係る情報提供方法は、第12の発明の構成に加えて、特徴ベクトル生成ステップ

は、単語特徴ベクトルの成分ごとに割当てられる第2の重みを算出する第2の重み算出ステップをさらに含み、加算ステップは、複数の履歴情報の各々について、履歴情報に含まれる単語に対応する単語特徴ベクトルの成分に第2の重みを付けた第2のベクトルを算出し、第2のベクトルに第1の重みを付けた第1のベクトルを算出し、算出された第1のベクトルを互いに加算するステップを含む。

【0033】第14の発明によると、第2の重み算出ステップは、単語特徴ベクトルの成分ごとに割当てられる第2の重みを算出する。加算ステップは、複数の履歴情報の各々について、履歴情報に含まれる単語に対応する単語特徴ベクトルの成分に第2の重みを付けた第2のベクトルを算出し、第2のベクトルに第1の重みを付けた第1のベクトルを算出し、算出された第1のベクトルを互いに加算する。これにより、利用者の履歴情報ごとの特徴ベクトルを算出する場合において、単語ベクトルの成分に対応させた第2の重み（たとえば、単語ベクトルの特定の成分の出現頻度に対応させて設定する重み）と、単語に対応させた第1の重み（たとえば、単語の出現頻度に対応させて設定する重み）とに基づいて履歴情報ごとの特徴ベクトルを算出できる。その結果、単語特徴記憶ステップにて記憶された単語特徴ベクトルを適切に更新することができる。

【0034】第15の発明に係る情報提供方法は、第9～14のいずれかの発明の構成に加えて、利用者を識別する利用者識別ステップをさらに含み、利用履歴記憶ステップは、利用者別に情報を記憶する複数の記憶ステップを含み、単語特徴更新ステップは、利用者識別ステップにより識別された利用者が変更されたことを検知すると、単語特徴ベクトルを、識別された利用者の履歴情報とその特徴ベクトルとに基づいて、単語特徴記憶ステップにて記憶した内容を更新するステップを含む。

【0035】第15の発明によると、利用者識別ステップは、利用者を識別し、利用履歴記憶ステップは、利用者別に情報を記憶する。単語特徴更新ステップは、利用者識別ステップにより識別された利用者が変更されたことを検知すると、単語特徴ベクトルを、識別された利用者の履歴情報とその特徴ベクトルとに基づいて、単語特徴記憶ステップにて記憶した内容を更新する。これにより、利用者ごとの履歴情報に基づいて、単語特徴記憶ステップにて記憶した内容を更新することができる。その結果、利用者ごとに利用者の利用履歴から利用者の興味の高い情報を迅速かつ的確に抽出し、その情報を用いて利用者の興味を反映した情報を検索し提供することができる情報提供方法を提供することができる。

【0036】第16の発明に係る情報提供方法は、第9～15のいずれかの発明の構成に加えて、提供可能情報特徴データは、各提供可能な情報に出現する単語に対応する単語特徴ベクトルを含み、利用履歴情報特徴データ

は、各分類ごとに含まれる利用履歴情報に出現する単語に対応する単語特徴ベクトルを含み、提供情報検索ステップは、提供可能情報特徴データに含まれる単語特徴ベクトルと利用履歴情報特徴データに含まれる単語特徴ベクトルとにより検索するステップを含む。

【0037】第16の発明によると、提供情報検索ステップは、提供可能情報特徴データに含まれる単語特徴ベクトルと利用履歴情報特徴データに含まれる単語特徴ベクトルとにより検索する。これにより、ベクトルによる演算を用いて、情報を検索することができ、利用者に提供する情報をさらに正確に検索できる情報提供方法を実現できる。その結果、利用者の利用履歴から利用者の興味の高い情報を迅速かつ的確に抽出し、その情報を用いて利用者の興味を反映した情報を検索し提供することができる情報提供方法を提供することができる。

【0038】第17の発明に係る機械読取可能な記録媒体は、コンピュータを用いて利用者に情報を提供する方法を実現するプログラムを記録した記録媒体であって、情報提供方法は、複数の単語と、各単語の特徴を表わす単語特徴ベクトルとを対応付けて記憶する単語特徴記憶ステップと、利用者に対して提供可能な情報と、提供可能な情報の各々に対して単語特徴記憶ステップにて記憶した単語および単語特徴ベクトルに基づいて生成された提供可能情報特徴データとを準備する提供可能情報準備ステップと、利用者が利用した情報の履歴情報を記憶する利用履歴情報記憶ステップと、単語特徴記憶ステップにて記憶した単語および単語特徴ベクトルを参照し、利用履歴情報記憶ステップにて記憶した履歴情報を、各履歴情報に含まれる単語および対応の単語特徴ベクトルに基づいて複数のカテゴリに分類する分類ステップと、複数のカテゴリはそれぞれ、単語および単語特徴ベクトルに関連して定められる利用履歴情報特徴データにより特徴付けられ、単語特徴記憶ステップにて記憶した単語および単語特徴ベクトルを参照し、複数の履歴情報の各々についての特徴ベクトルを生成する特徴ベクトル生成ステップと、予め定められた条件が満たされると、複数の履歴情報と其特征ベクトルとに基づいて、単語特徴記憶ステップにて記憶した内容を更新する単語特徴更新ステップと、提供可能情報準備ステップにて準備した提供可能な情報のうち、利用履歴情報特徴データと予め定められた関係にある提供可能情報特徴データを有するものを検索し出力する提供情報検索ステップとを含む。

【0039】第17の発明によると、単語特徴記憶ステップは、複数の単語と、各単語の特徴を表わす単語特徴ベクトルとを対応付けて記憶する。提供可能情報準備ステップは、利用者に対して提供可能な情報と、提供可能な情報の各々に対して単語特徴記憶ステップにて記憶した単語および単語特徴ベクトルに基づいて生成された提供可能情報特徴データとを準備する。利用履歴情報記

憶ステップは、利用者が利用した情報の履歴情報を記憶する。分類ステップは、単語特徴記憶ステップにて記憶した単語および単語特徴ベクトルを参照し、利用履歴情報記憶ステップにて記憶した履歴情報を、各履歴情報に含まれる単語および対応の単語特徴ベクトルに基づいて複数のカテゴリに分類する。特徴ベクトル生成ステップは、単語特徴記憶ステップにて記憶した単語および単語特徴ベクトルを参照し、複数の履歴情報の各々についての特徴ベクトルを生成する。単語特徴更新ステップは、予め定められた条件が満たされると、複数の履歴情報と其特征ベクトルとに基づいて、単語特徴記憶ステップにて記憶した内容を更新する。提供情報検索ステップは、提供可能情報準備ステップにて準備した提供可能な情報のうち、利用履歴情報特徴データと予め定められた関係にある提供可能情報特徴データを有するものを検索し出力する。これにより、単語特徴記憶ステップにより記憶される内容は逐次更新され、最新の利用者の履歴情報に基づく単語およびその単語の単語特徴ベクトルとが記憶される。これに加えて、利用履歴情報自体ではなく単語特徴記憶ステップにて記憶した内容に基づき分類されたカテゴリごとの特徴データに従って提供情報を検索することができるため、利用者に提供すべき情報を迅速かつ的確に検索できる。その結果、利用者の利用履歴から利用者の興味の高い情報を迅速かつ的確に抽出し、その情報を用いて利用者の興味を反映した情報を検索し提供することができる情報提供方法を実現するプログラムを記録した記録媒体を提供することができる。

【0040】第18の発明に係る機械読取可能な記録媒体は、第17の発明の構成に加えて、単語特徴更新ステップは、複数の履歴情報と其特征ベクトルとに基づいて、単語特徴記憶ステップにて記憶した内容に単語およびその単語に対応する単語特徴ベクトルを追加するステップを含む。

【0041】第18の発明によると、単語特徴更新ステップは、複数の履歴情報と其特征ベクトルとに基づいて、単語特徴記憶ステップにて記憶した内容に単語およびその単語に対応する単語特徴ベクトルを追加する。これにより、単語特徴記憶ステップにて記憶されていなかった単語およびその単語の単語特徴ベクトルを生成し、記憶させることができる。その結果、当初記憶させた単語およびその単語の単語特徴ベクトルに加えて、多数の単語に基づいて、利用者の利用履歴から利用者の興味の高い情報を迅速かつ的確に抽出し、その情報を用いて利用者の興味を反映した情報を検索し提供することができる情報提供方法を実現するプログラムを記録した記録媒体を提供することができる。

【0042】第19の発明に係る機械読取可能な記録媒体は、第17または18の発明の構成に加えて、単語特徴更新ステップは、複数の履歴情報と其特征ベクトルとに基づいて、単語特徴記憶ステップにて記憶した単語

特徴ベクトルを更新するステップを含む。

【0043】第19の発明によると、単語特徴更新ステップは、複数の履歴情報とその特徴ベクトルとに基づいて、単語特徴記憶ステップにて記憶した単語特徴ベクトルを更新する。これにより、利用履歴情報の変化に対応させて単語特徴ベクトルを生成し、単語特徴更新ステップにて更新することができる。その結果、利用履歴情報の変化に対応させて当初記憶させた単語特徴ベクトルを更新して、更新された単語特徴ベクトルに基づいて、利用者の利用履歴から利用者の興味の高い情報を迅速かつ的確に抽出し、その情報を用いて利用者の興味を反映した情報を検索し提供することができる情報提供方法を実現するプログラムを記録した記録媒体を提供することができる。

【0044】第20の発明に係る機械読取可能な記録媒体は、第17～19のいずれかの発明の構成に加えて、特徴ベクトル生成ステップは、履歴情報に含まれる各単語に割当てられる第1の重みを算出する第1の重み算出ステップと、複数の履歴情報の各々について、履歴情報に含まれる単語に対応する単語特徴ベクトルに第1の重みを付けた第1のベクトルを算出し、算出された第1のベクトルを互いに加算する加算ステップとを含む。

【0045】第20の発明によると、第1の重み算出ステップは、履歴情報に含まれる各単語に割当てられる第1の重みを算出する。加算ステップは、複数の履歴情報の各々について、履歴情報に含まれる単語に対応する単語特徴ベクトルに第1の重みを付けた第1のベクトルを算出し、算出された第1のベクトルを互いに加算する。これにより、利用者の履歴情報ごとの特徴ベクトルを算出する場合において、単語に対応させた第1の重み（たとえば、単語の出現頻度に対応させて設定する重み）に基づいて履歴情報ごとの特徴ベクトルを算出できる。その結果、単語特徴記憶ステップにて記憶した単語特徴ベクトルを適切に更新することができる。

【0046】第21の発明に係る機械読取可能な記録媒体は、第17～19のいずれかの発明の構成に加えて、特徴ベクトル生成ステップは、単語特徴ベクトルの成分ごとに割当てられる重みを算出する重み算出ステップと、複数の履歴情報の各々について、履歴情報に含まれる単語に対応する単語特徴ベクトルの成分に重みを付けたベクトルを算出し、算出されたベクトルを互いに加算する加算ステップとを含む。

【0047】第21の発明によると、重み算出ステップは、単語特徴ベクトルの成分ごとに割当てられる重みを算出する。加算ステップは、複数の履歴情報の各々について、履歴情報に含まれる単語に対応する単語特徴ベクトルの成分に重みを付けたベクトルを算出し、算出されたベクトルを互いに加算する。これにより、利用者の履歴情報ごとの特徴ベクトルを算出する場合において、単語ベクトルの成分に対応させた重み（たとえば、単語ベ

クトルの特定の成分の出現頻度に対応させて設定する重み）に基づいて履歴情報ごとの特徴ベクトルを算出できる。その結果、単語特徴記憶ステップにて記憶した単語特徴ベクトルを適切に更新することができる。

【0048】第22の発明に係る機械読取可能な記録媒体は、第20の発明の構成に加えて、特徴ベクトル生成ステップは、単語特徴ベクトルの成分ごとに割当てられる第2の重みを算出する第2の重み算出ステップをさらに含み、加算ステップは、複数の履歴情報の各々について、履歴情報に含まれる単語に対応する単語特徴ベクトルの成分に第2の重みを付けた第2のベクトルを算出し、第2のベクトルに第1の重みを付けた第1のベクトルを算出し、算出された第1のベクトルを互いに加算するステップを含む。

【0049】第22の発明によると、第2の重み算出ステップは、単語特徴ベクトルの成分ごとに割当てられる第2の重みを算出する。加算ステップは、複数の履歴情報の各々について、履歴情報に含まれる単語に対応する単語特徴ベクトルの成分に第2の重みを付けた第2のベクトルを算出し、第2のベクトルに第1の重みを付けた第1のベクトルを算出し、算出された第1のベクトルを互いに加算する。これにより、利用者の履歴情報ごとの特徴ベクトルを算出する場合において、単語ベクトルの成分に対応させた第2の重み（たとえば、単語ベクトルの特定の成分の出現頻度に対応させて設定する重み）と、単語に対応させた第1の重み（たとえば、単語の出現頻度に対応させて設定する重み）とに基づいて履歴情報ごとの特徴ベクトルを算出できる。その結果、単語特徴記憶ステップにて記憶された単語特徴ベクトルを適切に更新することができる。

【0050】第23の発明に係る機械読取可能な記録媒体は、第17～22のいずれかの発明の構成に加えて、利用者を識別する利用者識別ステップをさらに含み、利用履歴記憶ステップは、利用者別に情報を記憶する複数の記憶ステップを含み、単語特徴更新ステップは、利用者識別ステップにより識別された利用者が変更されたことを検知すると、単語特徴ベクトルを、識別された利用者の履歴情報とその特徴ベクトルとに基づいて、単語特徴記憶ステップにて記憶した内容を更新するステップを含む。

【0051】第23の発明によると、利用者識別ステップは、利用者を識別し、利用履歴記憶ステップは、利用者別に情報を記憶する。単語特徴更新ステップは、利用者識別ステップにより識別された利用者が変更されたことを検知すると、単語特徴ベクトルを、識別された利用者の履歴情報とその特徴ベクトルとに基づいて、単語特徴記憶ステップにて記憶した内容を更新する。これにより、利用者ごとの履歴情報に基づいて、単語特徴記憶ステップにて記憶した内容を更新することができる。その結果、利用者ごとに利用者の利用履歴から利用者の興味

の高い情報を迅速かつ的確に抽出し、その情報を用いて利用者の興味を反映した情報を検索し提供することができる情報提供方法を提供することができる。

【0052】第24の発明に係る機械読取可能な記録媒体は、第17～23のいずれかの発明の構成に加えて、提供可能情報特徴データは、各提供可能な情報に出現する単語に対応する単語特徴ベクトルを含み、利用履歴情報特徴データは、各分類ごとに含まれる利用履歴情報に出現する単語に対応する単語特徴ベクトルを含み、提供情報検索ステップは、提供可能情報特徴データに含まれる単語特徴ベクトルと利用履歴情報特徴データに含まれる単語特徴ベクトルとにより検索するステップを含む。

【0053】第24の発明によると、提供情報検索ステップは、提供可能情報特徴データに含まれる単語特徴ベクトルと利用履歴情報特徴データに含まれる単語特徴ベクトルとにより検索する。これにより、ベクトルによる演算を用いて、情報を検索することができ、利用者に提供する情報をさらに正確に検索できる情報提供方法を実現できる。その結果、利用者の利用履歴から利用者の興味の高い情報を迅速かつ的確に抽出し、その情報を用いて利用者の興味を反映した情報を検索し提供することができる情報提供方法を実現するプログラムを記録した記録媒体を提供することができる。

【0054】

【発明の実施の形態】以下、図面を参照しつつ、本発明の実施の形態について説明する。以下の説明では、同一の部品には同一の符号を付してある。それらの名称および機能も同じである。したがってそれらについての詳細な説明の繰返しは適宜省略する。

<第1の実施の形態>図1を参照して、本発明の第1の実施の形態に係る情報提供装置は、利用者が過去に利用した情報などを含む操作履歴データを入力する入力部2と、利用者を識別する利用者識別部4と、単語と単語の特徴を表わす単語ベクトルとをそれぞれ対応付けて記憶した単語辞書6と、利用者に提供可能な情報を記憶する提供可能データ記憶部8と、入力部2と利用者識別部4とに接続され、入力部2から入力された利用者の操作履歴データを利用者ごとに記憶する操作履歴データ記憶部10と、操作履歴データから生成した操作履歴ベクトルを記憶する操作履歴ベクトル記憶部12と、操作履歴ベクトルをクラスタリング処理して分類し、その分類ごとの特徴データである興味情報を記憶する興味情報記憶部14と、検索した結果をモニタ、プリンタなどに出力する出力部16と、利用者の興味が高い情報を検索する興味情報処理部30とを含む。

【0055】興味情報処理部30は、実際にはコンピュータ上で実行されるソフトウェアにより実現されるが、利用者を識別したり操作履歴データの種別を識別したりするデータ識別処理部32と、操作履歴データおよび提供可能データに含まれる単語の中で情報提供処理に不要

な単語を削除する不要語処理部34と、単語辞書6に記憶された単語に基づいて操作履歴データおよび提供可能データから単語を抽出する単語抽出処理部36と、単語辞書6に記憶された単語ベクトルおよび抽出された単語に基づいて操作履歴データごとの特徴を表わす操作履歴ベクトルを生成するベクトル生成処理部38と、操作履歴ベクトルを距離の近いベクトル同士に分類するクラスタリング処理部40と、所定のタイミングでクラスタリング処理部40に動作の指示を出力する更新処理部42と、クラスタリング処理部40で分類されたカテゴリごとに興味情報を生成する興味情報生成処理部44と、興味情報に基づいて提供可能データの中で利用者に対応した提供情報を検索する検索処理部46と、検索された提供情報を重要度順に並べ替える提示情報選定処理部48とを含む。

【0056】さらに、興味情報処理部30は、単語辞書6に記憶された単語ベクトルを更新したり、単語辞書に記憶されていない単語およびその単語に対応する単語ベクトルを単語辞書に記憶させたりする単語辞書修正処理部7と、所定のタイミングで単語辞書修正処理部7に動作の指示を出力する学習制御処理部37とを含む。

【0057】入力部2および利用者識別部4は、後述するコンピュータのキーボード、マウス、CD-ROM (Compact Disc Read-Only Memory) 駆動装置、FD (Flexible Disk) 駆動装置などであって、操作履歴データ、利用者識別コードを入力する。入力される操作履歴データは、たとえば、インターネットのホームページデータ、テレビ番組の情報であるEPG (Electronic Program Guide) データであって、利用者が同じホームページ、同じテレビ番組を一定時間以上見続けたり、あるホームページからデータをダウンロードしたり、あるテレビ番組を録画したりした場合のホームページデータ、EPGデータである。なお、入力された操作履歴データは、利用者識別コードとデータ種類識別コードとともに操作履歴データ記憶部10に記憶される。なお、このような入力部2を介さずにインターネットなどを操作した際の履歴データを固定ディスクに記憶しておき、コンピュータの内部処理により、この履歴データから操作履歴データを生成することもできる。

【0058】単語辞書6と提供可能データ記憶部8と操作履歴データ記憶部10と操作履歴ベクトル記憶部12と興味情報記憶部14とは、コンピュータの固定ディスク上の予め定めた記憶領域に存在し、該当するデータを記憶するものである。

【0059】以下の説明では、利用者にとって興味のあるテレビ番組の情報を、そのテレビ番組に対応するEPGデータで提供する装置について説明する。

【0060】図2を参照して、提供可能データ記憶部8に記憶されたデータは、放送予定の番組のEPGデータであって提供可能データの番号とそれぞれに対応したデ

ータ内容を含む。このEPGデータの例において、EPGデータは1行目に番組タイトル、2行目に放送日時、放送時間、チャンネル、いわゆるGコード（登録商標）、3行目に番組内容説明文を含む。

【0061】不要語処理部34は、操作履歴データおよび提供情報データに対して、利用者の興味には直接関係しない部分を不要語として以降の処理で無視する処理を行なう。本実施の形態では、図2および後述する図3に示すEPGデータのデータ内容の第2行目（放送日時、放送時間、チャンネル、Gコード）を不要語としている。なお、この不要語処理は、インターネット上のホームページデータを操作履歴データとした場合には、HTML（Hyper-Text Markup Language）のタグなどを取除くという処理になる。

【0062】図3を参照して、操作履歴データ記憶部10に記憶されたデータは、利用者が一定時間見続けた、録画したりした番組のEPGデータであって、操作履歴データの番号とそれぞれに対応したデータ内容を含む。データの内容については前述と同じであるため詳細な説明は繰返さない。

【0063】図4を参照して、単語辞書6は、単語とその単語に対応した単語ベクトルを含む。単語辞書6は、人手により単語およびその単語に対応する単語ベクトルが入力されて記憶される基本的な単語（以下、基本単語という。）と、後述する単語辞書修正処理により、基本単語に基づいて単語ベクトルが生成されて記憶される単語を含む。単語ベクトルは、文章中の単語が持つ概念と文脈との関係の程度を示したものであり、個々の単語が有する概念と予め定めた特徴単語との関係をベクトル形式にて表現したものである。n個の特徴単語を例として説明すると、ある単語の単語ベクトルはn次元ベクトルであって、その各要素の値は対応の特徴単語に対応して定められる。基本単語Jの単語ベクトルを $x_j = (x_j(1), \dots, x_j(m), \dots, x_j(n))$ とする。この各要素の値は、 $0 \leq x_j(m) \leq A$ （Aは正の定数、 $m=1 \sim n$ ）である。単語Jとm番目の特徴単語との間に全く関係がない場合には $x_j(m) = 0$ となり、関係がある場合には $x_j(m)$ はその関係の程度に応じた値を取る。以下の説明においては、 $A=1$ を用いて説明する。たとえば、単語ベクトルが10個の特徴単語（1）＝「娯楽・趣味」、特徴単語（2）＝「地名」、特徴単語（3）＝「自然」、特徴単語（4）＝「動的」、特徴単語（5）＝「感情」、特徴単語（6）＝「変化」、特徴単語（7）＝「複雑」、特徴単語（8）＝「喜楽」、特徴単語（9）＝「活動」、特徴単語（10）＝「明るさ」から成り立っているものとし、それぞれの要素の値が0か1かの2値である場合には、基本単語「お笑い」の単語ベクトルは（0, 0, 0, 0, 0, 0, 1, 1, 1）などと表わすことができる。図4に示す単語辞書は、10次元（ $n=10$ ）

の単語ベクトルを含む。

【0064】単語抽出処理部36は、図4に示す単語辞書6に記憶された単語により、図2に示す提供可能データおよび図3に示す操作履歴データのデータ内容から単語抽出処理を行なう。

【0065】学習制御処理部37は、入力部2から入力され操作履歴データ記憶部10に記憶された操作履歴データのデータ量が予め定められたデータ量を超えた場合、利用者識別部4から入力され操作履歴データ記憶部10に記憶された利用者の変更された場合、または前回の単語辞書修正指示出力から予め定められた時間を経過した場合などに、単語辞書修正処理部7へ単語辞書の修正処理の指示を出力する。この場合、操作履歴データから抽出された単語（単語辞書に記憶されていない単語）の中で、単語辞書6に記憶する単語を選択させることもできる。また、単語辞書6に記憶されている単語の中で、単語ベクトルを修正する単語を選択させることもできる。また、操作履歴データの中から操作履歴データの一部を選択して、選択された操作履歴データに基づいて単語辞書6の修正を行わせることもできる。

【0066】単語辞書修正処理部7は、たとえば、電子情報通信学会発行の信学技法A192-99（1993年1月22日発行）「大規模データベースからの連想検索」で提案されたブートストラップ学習に基づいて、単語辞書6の修正処理を行なう。このブートストラップ学習を用いて、操作履歴データに含まれる単語に基づいて、単語辞書6に記憶されていない単語とその単語に対応した単語ベクトルとを生成したり、単語辞書6に記憶されている単語の単語ベクトルを修正したりする。基本単語とその単語に対応する単語ベクトルは人手により単語辞書6に入力され、入力された基本単語に基づいて、ブートストラップ学習を行うことにより基本単語以外の単語の単語ベクトルが生成される。なお、この単語辞書修正処理については、このブートストラップ学習による処理には限定されない。また、この処理の詳細については、後述する。

【0067】ベクトル生成部38は、図4に示す単語辞書6に記憶された単語ベクトルにより、図3に示す操作履歴データごとに抽出された単語に対応した単語ベクトルのベクトル和を計算し、操作履歴ベクトルを生成する。たとえば、図3の操作履歴データ1の場合には、不要語処理部34でデータ内容の第2行目を不要語とする処理をして、不要語の処理がされたデータに対して単語抽出処理部36により図4に示す単語辞書6に記憶された単語を抽出する。この結果、図3の操作履歴データ1について「今日」、「旅行」、「気分」、「温泉」、「紹介」の5単語を抽出する。図4に示す単語辞書6に記憶された抽出単語の各々に対応する単語ベクトルの和を計算する。この結果、操作履歴データ1のベクトルは（3, 2, 1, 2, 1, 0, 0, 0, 2, 1）となる。



これを操作履歴ベクトル記憶部12に記憶する(図5)。

【0068】クラスタリング処理部40は、利用者の操作履歴である操作履歴データから生成した複数の操作履歴ベクトルに基づいて自己組織化マップを用いて、操作履歴データを後述するカテゴリに分類する。クラスタリング処理の手法は、前述の操作履歴ベクトルを互いに近い(「近い」とはベクトル間の互いのユークリッド距離または余弦角が小さいことを意味する。以下同じ。)もの同士に分類できるものであればよく、特定のものには限定されない。たとえば、多次元空間内に存在する個々の要素に対してクラスタリング処理を行なう手法は、階層的手法によるものや、非階層的手法によるものなどがある。

【0069】本実施の形態におけるクラスタリング手法は、自己組織化マップを用いる。自己組織化マップの学習モデルについての詳細は、T. Kohonen, "The Self-organizing map", Proc. IEEE, vol. 78(9), p. 1464~80に記載がある。自己組織化マップは、T. Kohonenによって提案された手法で、ニューラルネットワークに基づく教師なし競合学習モデルであり、多次元特性値で表現されるベクトルデータを、通常2次元に位相配置できる特徴がある。

【0070】ここでは、単語ベクトルで表現された各操作履歴ベクトルを学習用の入力データとして教師なし学習を行なう。出力層の各ユニットは2次元空間に位相配置されており、入力データと同次元のベクトルを有している。学習は、各ユニットのベクトルを入力データに選択的に近づけることによって進められる。まず、ある入力データに対して最も近いベクトルを持つユニットを見つける。次に、このユニットの近傍領域内にある複数のユニットについて、それぞれのベクトルを入力データに近づける。この近傍領域以外のユニットのベクトルは、更新せずそのままとする。一定回数の学習を行なった後、各操作履歴ベクトルを最も近いベクトルを持つユニットに写像することで終了する。この結果、単語ベクトルが同一と見なされる操作履歴データは、同一のユニットに、特徴が類似していると見なされる操作履歴データは、近接のユニットに配置されて分類される。この同一および近接のユニットについてまとめたものをカテゴリという。

【0071】更新処理部42は、クラスタリング処理部40に対して、クラスタリング処理の指示を与える。たとえば、更新処理部42は、一定時間間隔でクラスタリング処理を指示したり、操作履歴データ記憶部10に一定量以上操作履歴データが蓄積するとクラスタリング処理を指示したりする。

【0072】興味情報生成処理部44は、カテゴリに分類された操作履歴データに含まれる単語を調べて、分類ごとに出現する単語を抽出し興味情報とする。たとえ

ば、図5に示す操作履歴データ1~5を自己組織化マップによる手法でクラスタリング処理をすると、図6に示すようにカテゴリ1~3に分類される。分類されたカテゴリごとに出現単語をまとめると、図7のようなカテゴリごとの興味情報となる。これを興味情報記憶部14に記憶する。

【0073】検索処理部46は、図2に示す提供可能データと興味情報とから、利用者の興味に合致する提供情報を検索する。検索処理部46は、カテゴリごとに興味情報の単語と、提供可能データに含まれる単語との一致度(「検索点数1」という。)を以下の式1で評価する。

〔式1〕

検索点数1 = 提供可能データに含まれる興味情報の単語の数  $\times$  100 / 興味情報の単語の数

たとえば、図7に示すカテゴリ1における図2の提供可能データ1の検索点数1について説明する。提供可能データ1に出現する単語は、不要語処理部34にて不要語が処理され、単語抽出処理部36により抽出された「ニュース」、「テレビ」、「事件」、「スポーツ」の4単語である。これらは、カテゴリ1の興味情報である「ニュース」、「スポーツ」などの8単語の中の3単語(「ニュース」、「テレビ」、「スポーツ」)を含む。よって検索点数1は、 $3 \times 100 / 8 = 37.5$ となる。

【0074】提示情報選定処理部48は、検索処理部46における検索結果が、利用者の操作履歴データと重複していた場合には当該検索結果を除外したり、提供する情報を重要度順に並べ替えたりする。

【0075】前述のように興味情報処理部30は実際には、パーソナルコンピュータまたはワークステーションなど、コンピュータ上で実行されるソフトウェアにより実現される。図8に、利用者に対する情報提供装置の一例であるコンピュータの外観を示す。図8を参照して、このコンピュータ60は、FD駆動装置82およびCD-ROM駆動装置84とを備えたコンピュータ本体70と、モニタ88と、プリンタ86と、キーボード76と、マウス74とを含む。

【0076】図9に、このコンピュータ60の構成をブロック図形式で示す。図9に示すように、コンピュータ60はFD駆動装置82およびCD-ROM駆動装置84に加えて、相互にバスで接続されたCPU72(Central Processing Unit)と、メモリ78と、固定ディスク80とを含む。FD駆動装置82にはFD90が装着される。CD-ROM駆動装置84にはCD-ROM92が装着される。

【0077】既に述べたように、利用者に対して情報を提供する装置は、コンピュータハードウェアとCPU72により実行されるソフトウェアとにより実現される。一般的にこうしたソフトウェアは、FD90、CD-R



OM92などの記録媒体に格納されて流通し、FD駆動装置82またはCD-ROM駆動装置84などにより記録媒体から読取られて固定ディスク80に一旦格納される。さらに固定ディスク80からメモリ78に読出されて、CPU72により実行される。図8および図9に示したコンピュータのハードウェア自体は一般的なものである。したがって、本発明の最も本質的な部分は、FD90、CD-ROM92、固定ディスク80などの記録媒体に記録されたソフトウェアである。

【0078】なお、図8および図9に示したコンピュータ自体の動作は周知であるので、ここではその詳細な説明は繰返さない。

【0079】図10を参照して、操作履歴データ処理について説明する。ステップ2（以下、ステップをSと略す。）にて開始される操作履歴データ処理は、操作履歴データの蓄積量が一定量を超えたり、前回の操作履歴データ処理から一定時間を経過したりした場合に行なわれる。

【0080】S4にて、CPU72は、入力部2などから入力されて操作履歴データ記憶部10に記憶されている操作履歴データを、固定ディスク80から読出しメモリ78に格納する。このとき、操作履歴データ記憶部10に記憶されている利用者識別コード、データ識別コードをあわせて読出す。

【0081】S6にて、データ識別処理部32は、S4で読出されたデータの種別を判別する。この判断は、S2で読出されたデータ識別コードにより行なう。以下の説明では、データ識別処理部32がデータの種別をEPGデータであると判断したものとして説明する。

【0082】S8にて、不要語処理部34は、S4で読出された操作履歴データに対する不要語を抽出し、これ以降の処理で無視する不要語処理を行なう。S6にて、識別したデータの種別に基づいて不要語が決定される。たとえば、図3に示すEPGデータの場合には、利用者の興味と直接関係のない、第2行目の放送日、放送時間、放送局、Gコードが不要語として決定される。操作履歴データ1の場合、このS8における処理の結果、操作履歴データ1は「今日は旅行気分」、「温泉を紹介」というデータとなる。

【0083】S10にて、CPU72は、不要語処理がされた操作履歴データを操作履歴データ記憶部10に一旦格納する。

【0084】S12にて、CPU72は、固定ディスク80に記憶された単語辞書6の中で操作履歴データの種別に合致した単語辞書を読出した後、S10にて、CPU72は、操作履歴データ記憶部10から一旦格納した操作履歴データを読出す。

【0085】S14にて、単語抽出処理部36は、読出した操作履歴データに対し、操作履歴データごとに単語辞書6に基づく単語抽出処理を行なう。

【0086】S16にて、ベクトル生成処理部38は、操作履歴ごとの操作履歴ベクトルを、S14で操作履歴ごとに抽出した単語に対応した単語ベクトルのベクトル和として生成する。

【0087】S18にて、CPU72は、操作履歴ごとの操作履歴ベクトルを操作履歴ベクトル記憶部12に記憶する。たとえば、図3に示す操作履歴データでは、操作履歴データ1～6に対応する6つの操作履歴ベクトルが記憶される（図5）。

【0088】次に、図11を参照して、提供可能データ処理について説明する。S22にて、提供可能データ記憶部8に記憶されたデータが更新されると、提供可能データ処理が開始される。なお、この開始タイミングは、前述の操作履歴データ処理の開始タイミングと同期を取る必要はない。

【0089】S23にて、CPU72は、単語辞書修正処理部7による単語辞書6の修正が完了しているか否かを判断する。単語辞書6の修正が完了していると（S23にてYES）、処理はS24に移される。一方、単語辞書6の修正が完了していないと（S23にてNO）、処理はS23に戻され、単語辞書修正処理部7による単語辞書6の修正の完了を待つ。

【0090】S24にて、CPU72は、入力部2などから入力されて提供可能データ記憶部8に記憶されている提供可能データを読出す。このとき、メモリ78に保持されている利用者識別コードおよびデータ識別コードに従い、利用者識別コードに合致しかつデータ識別コードに合致する提供可能データを読出す。

【0091】S26にて、不要語処理部34は、S24にて読出した提供可能データに対する不要語を抽出し、以降の処理で無視する不要語処理を行なう。

【0092】S28における処理はS10における処理に、S30における処理はS12における処理に、S32における処理はS14における処理に各々相当するため、ここでは詳細な説明は繰返さない。このような処理の結果、たとえば、図2に示す提供可能データ1については、「ニュース」、「テレビ」、「事件」、「スポーツ」の4つの単語が抽出される。

【0093】S34にて、CPU72は、提供可能データごとに抽出された単語を提供可能データ記憶部8に記憶させる。

【0094】次に、図12を参照して、提供情報検索処理について説明する。S44にて、CPU72は、更新処理部42から提供情報検索処理の更新指示が出されているか否かを判断する。更新指示部42からの更新指示がされていると（S44でYES）、処理はS46へ移される。一方、更新指示がされていないと（S44でNO）、処理はS44に戻され、更新指示を待つ。

【0095】S46にて、CPU72は、操作履歴ベクトル記憶部12に記憶された操作履歴ごとの操作履歴ベ

クトルを讀出す(図5)。

【0096】S48にて、クラスタリング処理部40は、讀出された操作履歴ベクトルに自己組織化マップ手法を適用することにより、ベクトル同士の近さによって操作履歴データをカテゴリごとに分類する。図5に示す6個の操作履歴データ1から6は、図6に示す3つのカテゴリに分類される。カテゴリ1は、ニュースやスポーツに関し、カテゴリ2は旅行に関し、カテゴリ3はバラエティに関する。このカテゴリは操作履歴データの特徴をもとに分類されたものであって、それぞれのカテゴリが利用者の興味を示していると考えられることができる。

【0097】S50にて、興味情報生成処理部44は、S48にて分類されたカテゴリから利用者の興味情報を生成する。図4に示す単語辞書6に基づいて、各カテゴリに含まれる操作履歴データに含まれるすべての単語を利用者の興味情報とする。たとえば、図6のカテゴリ3の興味情報は「爆笑」、「コント」、「気分」の3つの単語で表わされる(図7)。

【0098】S52にて、興味情報生成処理部44は、S50にて生成した興味情報を興味情報記憶部14に記憶する。

【0099】S54にて、CPU72は、提供可能データ記憶部8から、提供可能データを読み出し、S56にて、CPU72は、興味情報記憶部14から興味情報データを読み出す。

【0100】S58にて、検索処理部46は、S56にて読み出した興味情報データに基づいて、S54にて読み出した提供可能データの中から利用者に対応した興味情報を検索する。図7に示す各カテゴリについて、図2に示す提供可能データのそれぞれについての検索点数1を算出する。それらの検索点数1の中で1位のものを示すと次のようになる。

【0101】

カテゴリ1	1位	提供可能データ1	37.5
カテゴリ2	1位	提供可能データ3	28.6
	1位	提供可能データ5	28.6
カテゴリ3	1位	提供可能データ2	33.3

S60にて、提示情報選定処理部48は、S58で検索した各カテゴリごとの1位の提供可能データの出力順序を定める。図6に示すように、カテゴリ1には3件の操作履歴データが、カテゴリ2には2件の操作履歴データが、カテゴリ3には1件の操作履歴データが含まれている。操作履歴データが多く含まれるカテゴリは、利用者の興味が集まっており重要度が高いと考えられる。そのため、カテゴリを操作履歴データ件数の多い順に並べる。並べたカテゴリの順に従って、各カテゴリにおける検索点数1の1位の提供可能データを出力する。図6に示す場合には、カテゴリ1、カテゴリ2、カテゴリ3の順になり、提供情報の出力順序は、提供可能データ1、提供可能データ3、提供可能データ5、提供可能データ

2の順となる。

【0102】S62にて、CPU72は、S60にて定めた順序に従って出力部16から利用者に提供可能データを出力する。

【0103】図13および図14を参照して、単語辞書修正処理について説明する。S72にて、CPU72は、学習制御処理部37から単語辞書修正処理部7へ単語辞書の学習制御の指示がされたか否かを判断する。学習制御の指示があると(S72にてYES)、処理はS74へ移される。一方、学習制御の指示がないと(S72にてNO)、処理はS72へ戻され、学習制御の指示を待つ。

【0104】S74にて、CPU72は、変数1を初期化(1=1)する。S76にて、CPU72は、操作履歴データ記憶部10から操作履歴データ(1)を読み出す。読み出した操作履歴データ(1)は、操作履歴データごとにメモリ78の所定のレコードに格納する。この場合、図3に示す操作履歴データ1のデータ内容がレコード1に格納される。

【0105】S78にて、CPU72は、所定のレコードに格納された操作履歴データ(1)から単語を抽出する。この場合、抽出される単語は、単語辞書6に単語およびその単語に対応する単語ベクトルが記憶された基本単語と、基本単語以外の単語とを含む。

【0106】S80にて、CPU72は、変数1について1=1+1の演算を行なう。S82にて、CPU72は、変数1が操作履歴データ数を上回っているか否かを判断する。変数1が操作履歴データ数を上回っていると(S82にてYES)、処理はS84へ移される。一方、変数1が操作履歴データ数以下であると(S82にてNO)、処理はS76へ戻され、操作履歴データ記憶部10からさらに別の操作履歴データを読み出して単語抽出処理を行なう。

【0107】S84にて、CPU72は、操作履歴データから抽出された全ての単語に基づいて重みを算出する。算出された特徴単語の重みと単語ベクトルの重みとは、メモリ78に格納される。なお、これら重みに関する詳細は後述する。

【0108】S86にて、CPU72は、変数1を初期化(1=1)する。S88にて、CPU72は、操作履歴データ(1)から抽出した単語に含まれる基本単語の単語ベクトルに重みを付けて、操作履歴データ(1)の特徴ベクトルRSV(1)を以下の式2により算出する。

【式2】

$$RSV(1) = c \sum \{ W(J) P(1, J) (a(1) \times j(1), \dots, a(n) \times j(n)) \}$$

この式2において、演算Σは、Jが1から操作履歴データ(1)に含まれる基本単語の数までについて行なわれる。cはベクトルRSV(1)を所定の長さにするため

の係数、 $W(J)$ は操作履歴データにおける基本単語 $J$ の重み、 $P(I, J)$ は利用履歴データ $(I)$ (レコード $I$ )における基本単語 $J$ の出現頻度、 $(a(1), \dots, a(n))$ はそれぞれの特徴単語の重み、 $(x_j(1), \dots, x_j(n))$ は基本単語 $J$ の単語ベクトルを表わしている。なお、 $x_j(1)$ は、基本単語 $J$ の特徴単語 $(1)$ に対するベクトルの成分であり、 $x_j(n)$ は、基本単語 $J$ の特徴単語 $(n)$ に対するベクトルの成分であり、いずれも0か1の値である。なお、本発明はこれに限定されるものではなく、特徴単語との関係の度合いを数値化したものを複数のビットデータで表現してもよい。

【0109】基本単語の重み $W(J)$ は、利用履歴データの中の基本単語 $J$ の出現頻度により算出される。たとえば、以下の式3により算出される。

【式3】

$$W(J) = \log(N/df(J))$$

この式3において、 $N$ は全レコード数、 $df(J)$ は、基本単語 $J$ を含むレコード数を示す。すなわち、あるレコードに偏って出現する基本単語ほど重要であるため、重み $W(J)$ が大きくなる。

【0110】特徴単語の重み $(a(1), \dots, a(n))$ は、基本単語中の各特徴単語の出現頻度により算出される。たとえば、以下の式4により算出される。

【式4】

$$a(m) = \log(NC/fc(m)), (m=1 \sim n)$$

この式4において、 $NC$ は基本単語の数、 $fc(m)$ は特徴単語 $(m)$ が付与された基本単語の数を示す。すなわち、多くの基本単語に付与されている特徴単語は単語を意味的に分類する上での効果が小さいため、特徴単語の重み $a(m)$ は小さくなる。たとえば、特徴単語 $(3)$ ＝「自然」、特徴単語 $(4)$ ＝「動的」であって、基本単語の中で、特徴単語「動的」が少ない基本単語にしか付与されておらず、かつ、特徴単語「自然」が多くの基本単語に付与されている場合、式4により算出される重みは、 $a(3) < a(4)$ となる。多くの基本単語に付与されていない特徴単語 $(4)$ の重み $a(4)$ の方が大きくなる。

【0111】このような処理は、 $n$ 次元の単語ベクトルを有効的に活用するために、 $n$ 個の特徴単語が均等に基本単語に付与されている必要があるために行なわれる。このため、上述の式2は、式3および式4を用いて、レコード中に含まれている基本単語について、ブートストラップ学習により特徴単語の偏りが小さくなるように、その単語ベクトルを特徴単語の重みで変調して、操作履歴データにおける単語の重みとレコード内の単語の頻度とにより重み付けして加算したベクトルの和を一定の長さに正規化したものである。なお、上述の特徴単語の重みと単語ベクトルの重みとは、S84にて算出されメモ

リ78に格納されている。

【0112】S90にて、CPU72は、S88にて算出したベクトル $RSV(I)$ をメモリ78に格納する。S92にて、CPU72は、変数 $I$ について $I = I + 1$ の演算を行なう。S94にて、CPU72は、変数 $I$ が操作履歴データ数を上回っているか否かを判断する。変数 $I$ が操作履歴データ数を上回っていると(S94にてYES)、処理は図14のS96へ移される。一方、変数 $I$ が操作履歴データ数以下であると(S94にてNO)、処理はS86へ戻され、さらに別のベクトル $RSV(I)$ を算出する。

【0113】S96にて、CPU72は、変数 $I$ を初期化( $I=1$ )する。S98にて、CPU72は、変数 $K$ を初期化( $K=1$ )する。S100にて、CPU72は、操作履歴データ $(I)$ から抽出した $K$ 番目の単語 $(I, K)$ を選択する。この場合、選択された単語 $(I, K)$ は、既に単語辞書6に単語ベクトルが登録されている基本単語であったり、単語辞書6に登録されていない単語であったりする。

【0114】S102にて、CPU72は、S100にて選択した単語 $(I, K)$ を含む操作履歴データの特徴ベクトル $RSV$ を抽出する。単語 $(I, K)$ を含む操作履歴データの特徴ベクトルに重みを付けて、重みを付けたものを加算する。これは、以下の式5で算出される。

【式5】

$$T'(I, K) = \sum \{P(I, K) RSV(I)\}$$

この式5において、演算 $\Sigma$ は、単語 $(I, K)$ を含む操作履歴データ $(I)$ の操作履歴データ $RSV(I)$ について行なわれる。すなわち、重み $P(I, K)$ が乗算されて加算されるのは、単語 $(I, K)$ を含む操作履歴データ $(I)$ の操作履歴データ $RSV(I)$ である。 $P(I, K)$ は、利用履歴データ $(I)$ (レコード $I$ )における単語 $(I, K)$ の出現頻度を示す。また、加算されて算出された単語 $(I, K)$ に対する $T'(I, K)$ は、メモリ78に単語 $(I, K)$ とともに格納される。

【0115】S104にて、CPU72は、変数 $K$ について $K = K + 1$ の演算を行なう。S106にて、CPU72は、変数 $K$ が操作履歴データ $(I)$ から抽出された単語数を上回っているか否かを判断する。変数 $K$ が操作履歴データ数 $(I)$ から抽出された単語数を上回っていると(S106にてYES)、処理はS108へ移される。一方、変数 $K$ が操作履歴データ $(I)$ から抽出された単語数以下であると(S106にてNO)、処理はS100へ戻され、操作履歴データ $(I)$ から抽出された他の単語について、 $T'(I, K)$ を算出する。なお、単語 $(I, K)$ は、同じ操作履歴データの中においても、異なる操作履歴データ間においても重複して出現することがある。このような場合、一旦算出した $T'$ と同じ単語について $T'$ を算出する場合には、再計算することなくS102の処理を行わないようにして、メモリ7

8に格納された $T'$ を用いると、重複する計算を省略できる。

【0116】S108にて、CPU72は、変数 $I$ について $I = I + 1$ の演算を行なう。S110にて、CPU72は、変数 $I$ が操作履歴データ数を上回っているか否かを判断する。変数 $I$ が操作履歴データ数を上回っていると(S110にてYES)、処理はS112へ移される。一方、変数 $I$ が操作履歴データ数以下であると(S110にてNO)、処理はS98へ戻され、次の操作履歴データにおける単語を抽出して、抽出された他の単語について、 $T'(I, K)$ を算出する。

【0117】S112にて、CPU72は、単語( $T, K$ )の単語ベクトルを変更する。この場合、単語( $I, K$ )が既に単語辞書6に単語およびその単語の単語ベクトルが記憶されている単語である場合には、 $T'(I, K)$ を正規化したものを単語ベクトルとして更新する。また、単語( $I, K$ )が単語辞書6に記憶されていない単語である場合には、単語( $T, K$ )と $T'(I, K)$ を正規化したものとを記憶する。

【0118】以上のような構造およびフローチャートに基づく興味情報提供装置の動作について説明する。

【0119】利用者は、放送予定のEPGデータが記憶されたFD90、CD-ROM92などをコンピュータのFD駆動装置82、CD-ROM駆動装置84に挿入し、EPGデータを提供可能データ記憶部8に記憶させる。また、利用者が一定時間見続けたり録画したりした番組のEPGデータを同様の方法により操作履歴データ記憶部10に記憶させる。このとき利用者は利用者識別コードを利用者識別部4から入力する。

【0120】入力された操作履歴データの種類のデータ識別処理部32により識別される(S6)。その種類に従って不要語処理部34により不要な単語の処理がなされる(S8)。その後、操作履歴データはユーザ識別コードとともに操作履歴データ記憶部10に記憶される(S10)。

【0121】入力された提供可能データに対しては、不要語処理部34により不要な単語の処理がなされ(S26)、単語抽出処理部36により単語辞書6に基づいてデータ内の単語が抽出され(S32)、提供可能データ記憶部8に記憶される(S34)。

【0122】記憶された操作履歴データに対しては、単語抽出処理部36により単語辞書6に基づいてデータ内の単語が抽出される(S14)。ベクトル生成処理部38により操作履歴データごとに抽出された単語に対応した単語ベクトルのベクトル和が計算される(S16)。その後、このベクトル和は操作履歴ベクトル記憶部12に記憶される(S18)。

【0123】前回の興味情報提供時から一定時間が経過していたり、操作履歴データが一定量以上蓄積していたりすると(S44にてYES)、操作履歴ベクトルが読

出されクラスタリング処理部40によりカテゴリに分類される(S48)。興味情報生成処理部44により、分類されたカテゴリごとに含まれる操作履歴データに出現するすべての単語を抽出した興味情報が生成される(S50)。検索処理部46により、利用者ごとの興味情報に基づいて、提供可能データの中で、利用者の興味が高い情報を検索して(S58)、提示情報選定処理部48により情報の出力順序を定める(S60)。出力部16は、その出力順序に従って、モニタ88やプリンタ86に利用者の興味が高い情報を出力する(S62)。

【0124】操作履歴データ記憶部10に記憶された操作履歴データのデータ量が予め定められたデータ量を超えた場合、利用者識別部4が利用者の変更を検知した場合、または、前回の単語辞書修正から予め定められた時間を経過した場合、単語辞書修正処理が開始される(S72にてYES)。操作履歴データ記憶部10に記憶された操作履歴データが読み出され(S76)、操作履歴データから基本単語が抽出される(S78)。操作履歴データの全てを読み出して、単語の抽出が終わると(S82にてYES)、抽出された基本単語などから、式3と式4とに基づいて、基本単語の重みと特徴単語の重みとが算出される(S84)。

【0125】操作履歴データごとに、式2に基づいて、基本単語の単語ベクトルと基本単語の重みと特徴単語の重みとを用いて、操作履歴データ(1)の特徴ベクトル $RSV(1)$ が算出される(S88)。操作履歴データの全てについて特徴ベクトルの算出が終わると(S94にてYES)、操作履歴データの中から単語(基本単語および基本単語以外の単語)を選択する(S100)。選択された単語について、その単語を含む操作履歴データの特徴ベクトルのベクトル和が算出される(S102)。全ての操作履歴データに含まれる全ての単語について特徴ベクトルのベクトル和の算出が終わると(S106にてYES、S110にてYES)、特徴ベクトルのベクトル和を正規化して、単語辞書6に反映させる(S112)。

【0126】図15を参照して、たとえば、操作履歴データ(1)がレコード1に格納され、抽出された基本単語は、「今日」(対応する単語ベクトルは $x(5)$ )、「旅行」(対応する単語ベクトルは $x(25)$ )、「気分」(対応する単語ベクトルは $x(4)$ )、「温泉」(対応する単語ベクトルは $x(2)$ )、「紹介」(対応する単語ベクトルは $x(15)$ )である。このような基本単語を含む操作履歴データ(1)の特徴ベクトル $RSV(1)$ が、式2に基づいて、基本単語の単語ベクトルと基本単語の重みと特徴単語の重みとを用いて算出される。このような操作履歴データごとの特徴ベクトルが算出される( $RSV(1), \dots, RSV(n), RSV(n+1), \dots$ )。図15に示すような操作履歴データおよび特徴ベクトルが格納されている場合において、基

本単語でない単語「露天」の単語ベクトルを算出するには、「露天」を含む操作履歴データの特徴ベクトルを加算する。すなわち、特徴ベクトル $RSV(n)$ と $RSV(n+1)$ とを重みを付けて加算して、特徴ベクトルのベクトル和を正規化して、「露天」の単語ベクトルとする。これを単語辞書6に反映させる。

【0127】なお、全ての操作履歴データに含まれる全ての単語について単語辞書の更新処理を行うのではなく、予め定められた条件に従って、操作履歴データに含まれる全ての単語から一部の単語を選択して単語辞書の更新処理を行うようにしてもよい。

【0128】以下、第1の実施の形態の変形例を示す。変形例は、図1における入力部2を受信部18に、出力部16を送信部22に変更したものである。受信部18と送信部22は、通信回線を介して他のコンピュータ、ワークステーションなどとネットワーク接続されており、それらとデータの送受信が可能である。

【0129】受信部18は利用者の操作履歴データを、他のコンピュータから利用者識別コードとともに受信する。

【0130】送信部22は、利用者識別コードに従い提示情報選定処理部48で選定した情報を受信データの発信元である他のコンピュータへ送信する。

【0131】たとえば、本発明の情報提供装置を、インターネットを接続するためのプロキシサーバ(proxy server)として機能するコンピュータにて実現させることができる。この場合、受信部18でクライアントパソコンからの操作履歴データを利用者識別コードとともに受信し、操作履歴データを利用者ごとに操作履歴データ記憶部12に記憶させる。興味情報処理部30で、この操作履歴データの特徴に応じて、利用者に提供する情報を検索し、送信部で提供する情報をクライアントパソコンに送信させる。

【0132】以上のようにして、本実施の形態に係る情報提供装置は、利用者の操作履歴をその特徴に応じたカテゴリに分類することにより、利用者の興味に関する情報を効率よく取得できる。その結果、利用者の興味に応じた情報を提供することができる。また、本実施の形態に係る情報提供装置をネットワーク上のサーバにて実現させることにより、クライアントパソコン各々に情報提供装置を有していなくても利用者の興味に応じた情報を提供することができる。

<第2の実施の形態>次に、本発明の第2の実施の形態に係る情報提供装置について説明する。

【0133】図16を参照して、本発明の実施の形態に係る情報提供装置は、図1に示す装置に加えて提供可能データベクトル記憶部20をさらに含む。また、情報提供装置は、図1に示す興味情報生成部44と検索処理部46とに代えて、それぞれこれらと多少機能の異なる興味情報生成処理部50と検索処理部52とを含む。

【0134】提供可能データベクトル記憶部20は、操作履歴ベクトル記憶部12と同様、コンピュータの固定ディスク上の予め定めた記憶領域に存在し、該当するデータを記憶する。提供可能ベクトルは、データ識別処理部32、不要語処理部34、単語抽出処理部36、ベクトル生成処理部38により操作履歴ベクトルと同様の処理により生成されるため、同じ処理についての詳細な説明は繰返さない。このような処理により、たとえば、図2に示す提供可能データ1~5から、図17に示す提供可能ベクトルが生成され、提供可能データベクトル記憶部20に記憶される。本実施の形態においては、各提供可能ベクトルの大きさが10なるように正規化処理がされる。

【0135】興味情報生成処理部50は、クラスタリング処理部40により分類された操作履歴データに基づいて、興味情報として興味ベクトルを生成する。興味ベクトルは、各カテゴリに属する操作履歴ベクトルのベクトル和を求め、さらに、ベクトルの大きさを10とする正規化処理を行なうことにより求める。たとえば、図6に示すカテゴリ2の興味ベクトルは、カテゴリ2に属する操作履歴データ1と操作履歴データ3の操作履歴ベクトルとの和となるため、図5の各々の操作履歴データのベクトル和を計算し(6, 5, 5, 2, 1, 0, 0, 0, 5, 2)となる。これをベクトルの大きさが10となるように正規化処理を行なうことにより、図18に示すように(5.5, 4.6, 4.6, 1.8, 0.9, 0, 0, 0, 4.6, 1.8)となる(小数点2位以下四捨五入)。これがカテゴリ2の興味ベクトルとなる。各カテゴリごとに求めた興味ベクトルを興味情報として興味情報記憶部14に記憶する。

【0136】検索処理部52は、図17に示す提供可能ベクトルと図18に示す興味ベクトルとから利用者の興味が強い情報を検索する。検索処理部52は、各カテゴリごとの興味ベクトル $V_q$ に基づいて提供可能ベクトル $V_i$ の評価度(「検索点数2」という。)を以下の式6で算出する。

【式6】

$$\text{検索点数2} = V_q \cdot V_i$$

この式6において、演算子「 $\cdot$ 」はベクトルの内積を表わしている。 $V_q$ 、 $V_i$ ともにその大きさを10とする正規化処理を行なっているため、検索点数2の値は0から100までの間の値を取り、ベクトルの距離が近いほど大きな値となる。

【0137】本実施の形態に係る情報提供装置は、以下のように動作する。なお、以下では第1の実施の形態と同様の動作についての詳細な説明は繰返さず、相違する動作についてのみ説明する。

【0138】図12のS54にて、CPU72は、提供可能データベクトル記憶部20から提供可能ベクトルを提供可能データとして読出し、S56にて、CPU72

は、興味情報記憶部14から興味ベクトルを興味情報データとして読出す。

【0139】S58にて、検索処理部52は、S54で読出した提供可能ベクトルの中で、S56で読出した興味ベクトルに基づいて、利用者の興味に応じた情報を検索する。図18に示す各カテゴリについて、図17に示す提供可能データのそれぞれについての提供可能ベクトルの検索点数2を算出する。それらの検索点数2の中で、1位のものを示すと次のようになる。

【0140】

カテゴリ1 1位 提供可能データ1 95  
 カテゴリ2 1位 提供可能データ3 85.6  
 カテゴリ3 1位 提供可能データ2 79.5  
 S60にて、提示情報選定処理部48は、S58にて検索した各カテゴリごとの1位の提供可能データの出力順序を定める。本実施の形態では提供可能データ1、提供可能データ3、提供可能データ2の順となる。

【0141】以上のように、本実施の形態に係る情報提供装置は、利用者の操作履歴データをその特徴に応じてカテゴリに分類し、分類ごとの興味ベクトルを生成する。また、提供可能データの単語ベクトルを生成する。それらのベクトル演算により利用者の興味に関する情報を精度よく取得できる。

【0142】今回開示された実施の形態はすべての点で例示であって制限的なものではないと考えられるべきである。本発明の範囲は上記した説明ではなくて特許請求の範囲によって示され、特許請求の範囲と均等の意味および範囲内でのすべての変更が含まれることが意図される。

【図面の簡単な説明】

【図1】 本発明の第1の実施の形態に係る情報提供装置のブロック図である。

【図2】 提供可能データの内容を示す図である。

【図3】 操作履歴データの内容を示す図である。

【図4】 単語辞書の内容を示す図である。

【図5】 操作履歴ベクトルの内容を示す図である。

【図6】 クラスタリング処理がなされたカテゴリの内容を示す図である。

【図7】 カテゴリ内の単語を示す図である。

【図8】 図9に示すコンピュータの外観図である。

【図9】 本発明の第1の実施の形態に係る情報提供装置を実現するコンピュータのブロック図である。

【図10】 操作履歴データ処理の概略を示すフローチャートである。

【図11】 提供可能データ処理の概略を示すフローチャートである。

【図12】 提供情報検索処理の概略を示すフローチャートである。

【図13】 単語辞書修正処理の概略を示すフローチャート（その1）である。

【図14】 単語辞書修正処理の概略を示すフローチャート（その2）である。

【図15】 操作履歴データごとの特徴ベクトルの内容を示す図である。

【図16】 本発明の第2の実施の形態に係る情報提供装置のブロック図である。

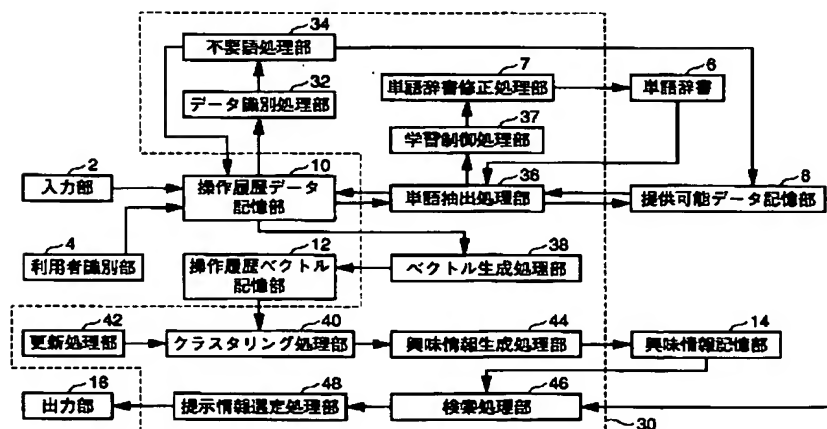
【図17】 提供情報ベクトルの内容を示す図である。

【図18】 カテゴリ内のベクトルを示す図である。

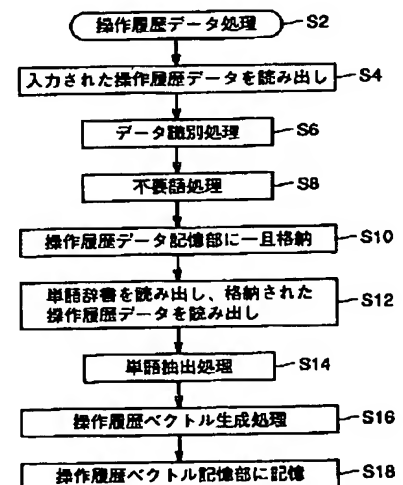
【符号の説明】

6 単語辞書、7 単語辞書修正処理部、8 提供可能データ記憶部、10 操作履歴データ記憶部、30 興味情報処理部、36 単語抽出処理部、37 学習制御処理部、38 ベクトル生成処理部、40 クラスタリング処理部、46 検索処理部

【図1】



【図10】





【図2】

提供可能データ	データ内容
提供可能データ1	ニューステレビ 1999/05/27 09:00~09:30 A放送局 Gコード:29395 事件とスポーツ
提供可能データ2	漫才 1999/05/27 10:00~10:30 B放送局 Gコード:03959 爆笑お笑い劇場
提供可能データ3	名所情報 1999/05/27 10:00~10:30 C放送局 Gコード:01982 名所と温泉を紹介
提供可能データ4	サスペンス 1999/05/27 21:00~22:50 A放送局 Gコード:190132 京都ミステリー
提供可能データ5	今日の経済 1999/05/27 11:00~11:30 D放送局 Gコード:60321 株価情報

【図4】

番号	単語	単語ベクトル(n=10)
1	お笑い	x1=(0,0,0,0,0,0,0,1,1,1)
2	温泉	x2=(1,1,1,0,0,0,0,0,1,0)
3	株価	x3=(0,0,0,0,0,1,1,0,0,0)
4	気分	x4=(0,0,0,1,1,0,0,0,0,0)
5	今日	x5=(0,0,0,1,0,0,0,0,0,1)
6	経済	x6=(0,0,0,0,0,1,0,1,1,0)
7	京都	x7=(0,1,1,0,0,0,1,0,0,0)
8	go	x8=(0,0,1,0,0,0,0,0,0,0)
9	コント	x9=(0,0,0,0,0,0,0,1,0,1)
10	劇場	x10=(0,1,0,0,0,0,0,1,0,1)
11	サスペンス	x11=(0,0,1,0,0,0,1,0,0,0)
12	事件	x12=(0,0,1,0,0,0,0,1,0,1)
13	紹介	x13=(1,0,0,0,0,0,0,0,0,0)
14	情報	x14=(0,1,0,0,0,0,0,0,0,0)
15	スポーツ	x15=(0,0,0,1,0,1,1,0,0,1)
16	テレビ	x16=(0,0,0,1,0,0,0,0,0,1)
17	ニュース	x17=(0,0,0,1,0,1,1,0,0,0)
18	NEWS	x18=(0,0,0,1,0,1,1,0,0,0)
19	爆笑	x19=(0,0,0,0,0,0,0,1,0,0)
20	プロ野球	x20=(0,0,0,1,0,1,1,0,0,0)
21	漫才	x21=(0,0,0,0,0,0,0,1,1,1)
22	名所	x22=(0,1,1,0,0,0,0,0,1,0)
23	ミステリー	x23=(0,1,1,0,0,0,1,0,0,0)
24	旅館	x24=(1,1,1,0,0,0,0,0,1,1)
25	旅行	x25=(1,1,0,0,0,0,0,0,1,0)

【図6】

カテゴリ	カテゴリ内の操作履歴データ
カテゴリ1	操作履歴データ2、操作履歴データ5、操作履歴データ6
カテゴリ2	操作履歴データ1、操作履歴データ3
カテゴリ3	操作履歴データ4

【図3】

操作履歴データ	データ内容
操作履歴データ1	今日は旅行気分 1999/05/22 09:00~09:30 A放送局 Gコード:12356 温泉を紹介
操作履歴データ2	今日のニュース 1999/05/22 19:00~19:30 B放送局 Gコード:53864 ニュースと株価とスポーツ
操作履歴データ3	旅行にgo! 1999/05/23 19:00~19:30 C放送局 Gコード:32185 温泉と旅館情報
操作履歴データ4	爆笑気分 1999/05/25 20:00~20:50 A放送局 Gコード:14564 コントで爆笑
操作履歴データ5	NEWS 1999/05/25 23:00~23:30 D放送局 Gコード:23461 ニュースと株価
操作履歴データ6	スポーツテレビ 1999/05/26 23:00~23:30 C放送局 Gコード:11034 プロ野球情報

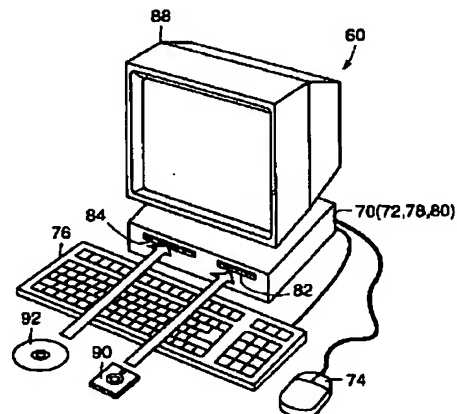
【図5】

操作履歴データ	操作履歴ベクトル
操作履歴データ1	(3, 2, 1, 2, 1, 0, 0, 0, 2, 1)
操作履歴データ2	(0, 1, 0, 3, 0, 3, 4, 1, 1, 2)
操作履歴データ3	(3, 3, 4, 0, 0, 0, 0, 0, 3, 1)
操作履歴データ4	(0, 0, 1, 1, 0, 0, 0, 3, 0, 3)
操作履歴データ5	(0, 0, 0, 2, 0, 3, 3, 0, 0, 0)
操作履歴データ6	(0, 0, 0, 3, 0, 2, 2, 0, 0, 2)

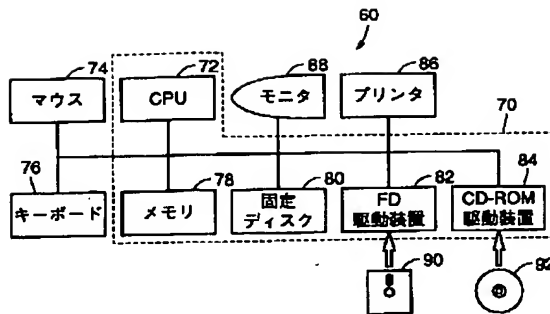
【図7】

カテゴリ	興味情報 (出現単語)
カテゴリ1	ニュース、スポーツ、今日、株価、NEWS、テレビ、プロ野球、情報
カテゴリ2	旅行、温泉、今日、気分、go、旅館、情報
カテゴリ3	爆笑、コント、気分

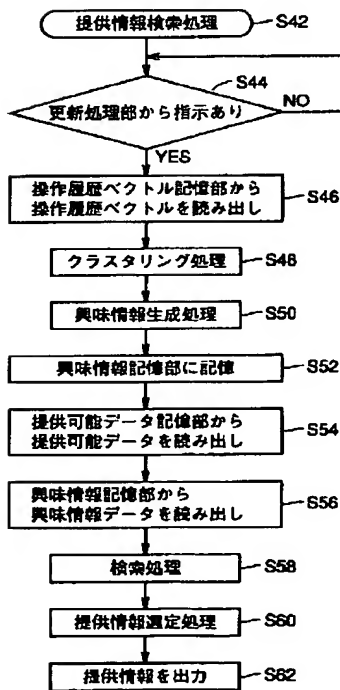
【図8】



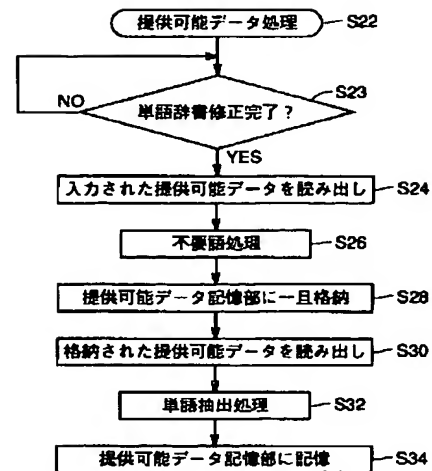
【図9】



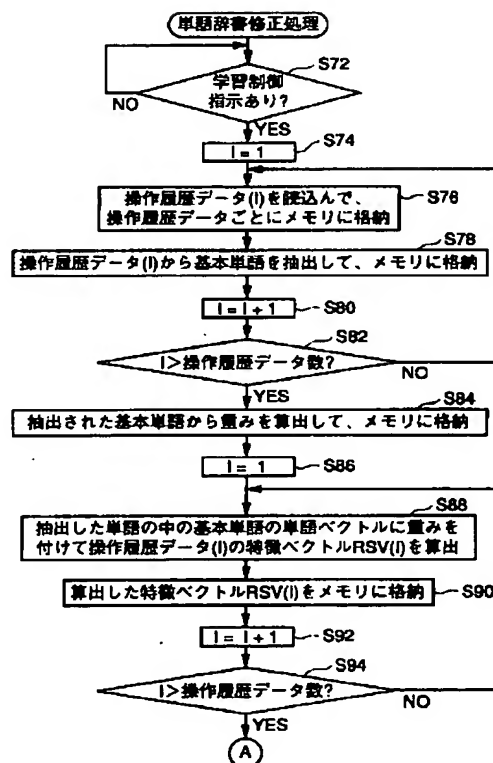
【図12】



【図11】



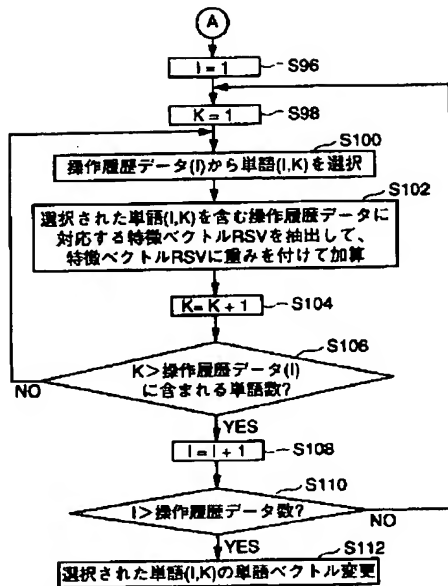
【図13】



【図18】

カテゴリ	興味ベクトル
カテゴリ1	(0, 0.7, 0, 5.3, 0, 5.3, 5.9, 0.7, 0.7, 2.6)
カテゴリ2	(5.5, 4.8, 4.8, 1.8, 0.9, 0, 0, 0, 4.6, 1.8)
カテゴリ3	(0, 0, 2.9, 2.9, 0, 0, 0, 8.7, 0, 2.9)

【図14】



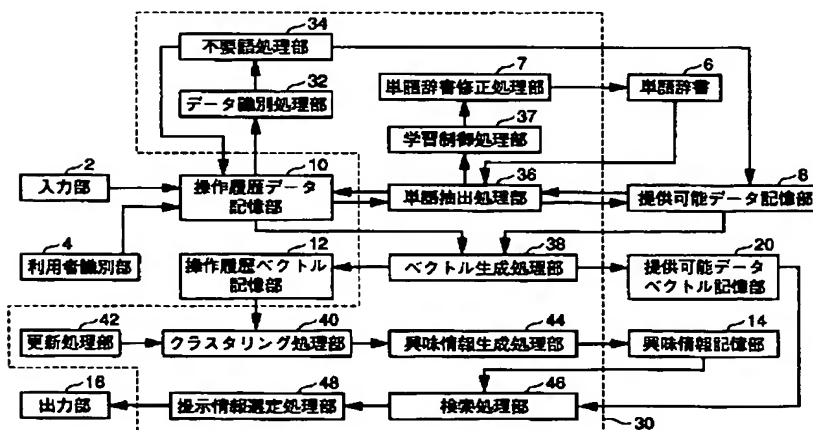
【図17】

提供可能データ	提供可能ベクトル
提供可能データ1	(0, 0, 2, 6, 0, 4, 6, 0, 2, 2)
提供可能データ2	(0, 1.8, 0, 0, 0, 0, 0, 7.3, 3.7, 5.5)
提供可能データ3	(2.6, 7.7, 5.2, 0, 0, 0, 0, 2.6, 1.5)
提供可能データ4	(0, 4.1, 6.3, 0, 0, 0, 6.3, 0, 2.1, 0)
提供可能データ5	(0, 3.2, 0, 3.2, 0, 6.3, 3.2, 3.2, 3.2, 3.2)

【図15】

操作履歴	レコード	単語 (単語ベクトル)						特徴ベクトル
データ1	1	今日(x(5))	旅行(x(25))	気分(x(4))	温泉(x(2))	紹介(x(15))	—	RSV(1)
データ2	2	今日(x(5))	ニュース(x(17))	ニュース(x(17))	株価(x(3))	スポーツ(x(15))	—	RSV(2)
データ3	3	旅行(x(25))	go(x(8))	温泉(x(2))	旅館(x(24))	情報(x(14))	—	RSV(3)
データ4	4	爆笑(x(19))	気分(x(4))	コント(x(9))	情報(x(14))	—	—	RSV(4)
...	...	...	...	...	...	...	...	...
データn	n	旅行(x(25))	露天(—)	温泉(x(2))	旅館(x(24))	情報(x(14))	気分(x(4))	RSV(n)
データn+1	n+1	名所(x(22))	紹介(x(13))	京都(x(7))	露天(—)	事件(x(12))	...	RSV(n+1)
...	...	...	...	...	...	...	...	...

【図16】



フロントページの続き

(72)発明者 芥子 育雄

大阪府大阪市阿倍野区長池町22番22号 シ  
ャープ株式会社内

Fターム(参考) 5B075 NR12 PP02 PP03 PQ02 PQ03  
PQ75 PR03 PR06 QM06